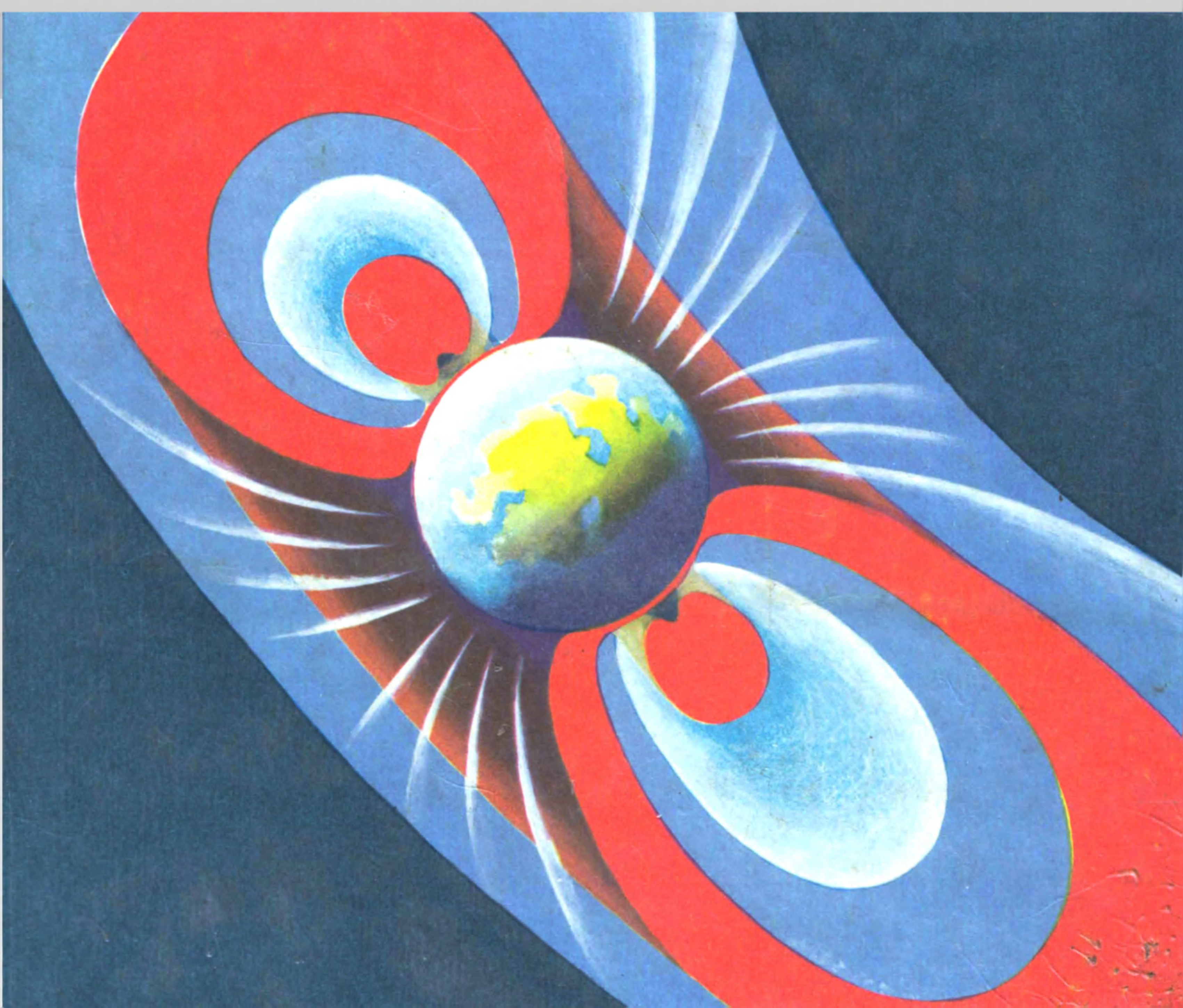
ESSAYS ABOUT

THE UNIVERSE

Boris A. Vorontsov-Vel'yaminov

Mir Publishers Moscow



ESSAYS ABOUT THE UNIVERSE

Boris A.Vorontsov-Vel'yaminov

ESSAYS ABOUT THE UNIVERSE

Translated from the Russian
by Alexander Repyev

Mir Publishers Moscow

Б. А. Воронцов-Вельяминов
ОЧЕРКИ О ВСЕЛЕННОЙ

Издательство «Наука» Москва

First published 1985
Translated from the revised 1980 Russian edition

На английском языке

Contents (Part 1)

PART 1

Introduction

The Starry Skies

A Heavenly Menagerie

Animals in the sky – a legacy of the past – are a frequent source of jokes about modern astronomers who had nothing to do with it. Who has not heard of the Big Bear and Little Bear in the sky? In the past high-ranking "wits" would send people from St. Petersburg to find out at Pulkovo Observatory "whether the astronomers had not forgotten to feed the Big Bear for the night...".

Thousands of years ago the fantasy of the ancients populated the sky with mythical creatures and animals, of which many do not exist in any zoo on earth. In the heavenly menagerie besides the Giraffe, Lion, Little Fox, Swan, Eagle, and many others there are the Unicorn, Dragon, Hydra, and so forth. These names have been preserved down to

the present time through tradition. They are continually mentioned in the scientific literature, enabling us to compare ancient and modern descriptions and observations of the sky. The ear of an expert astronomer is so accustomed to them that he fails to notice the incongruity of these names among the terms of modern science such as: integrals, spectrographs, milligrammes, thermoelements, and so on. Only the southern sky, studied more recently, is "populated" with the Air Pump, Microscope, Telescope, and other objects more contemporary to us.

Many colourful legends of ancient times, notably those of Greece and Rome, found their image in the sky. This is one of them. The legend is quite long, we'll tell it in abridged form.

"Once upon a time there lived a king called Ptolemy. He had a wife called Veronica. Once the king went to war, but things became very difficult for him. The queen was very worried and prayed to the goddess Venus. The queen promised the goddess to bring her own hair to the altar as a gift. Let the king conquer, and the queen would bring the sacrifice. Runners brought news of the victory and the queen's braids lay on the altar of the temple. They did not know how to make wigs then and so the queen remained without braids. The king returned in triumph, but when he saw the queen shorn of her lovely hair he was very upset. He became sad until Kommon, the king's astronomer, told him: "Do not grieve, king, do not be sad. Look at the sky, do you see the small stars there in the dark sky? These are your Veronica's hair, shining up in the sky."

The legend does not tell us whether the king was soothed by this news, but it informs us how the astronomer Kommon discovered the constellation "Veronica's Hair". The idea seems strange to us now – the ascension of "hair" into the sky. Some women of our days would ask in perplexity: "What kind of sacrifice is this, to go without braids!" Thus throughout the ages views have changed on the inviolability of hair.

And here is another legend about the stars that is known in different versions. We shall tell it in our style, to our taste.

"On the shores of a southern blue sea, lay an ancient, rocky land called Ethiopia. Long-long ago its ruler was King Cepheus. He had a wife, named Cassiopeia, and an only daughter, the beautiful princess Andromeda. When Andromeda grew up, there was no Ethiopian girl more beautiful than the princess.

"Cassiopeia was proud of the beauty of her daughter. So she began to brag in front of the whole world, and compared her daughter's beauty to the beauty of the goddesses.

"The gods then became angry and they sent a terrible calamity to Ethiopia. Every day a terrible monster Whale rose out of the sea and threatened to destroy the country.

"In order to appease the insatiable Whale each day the people gave him a young girl to devour, as the monster demanded. Soon there were no more young girls in the poor country, and Cepheus prayed to the gods and asked them to remove the terrible punishment from the country.

"The gods answered Cepheus, 'Your request will be honoured, but give the Whale as a sacrifice your only, beloved daughter – throw Andromeda, the beautiful princess, to be eaten by it.'

"Cepheus sobbed a long time and the queen sobbed a long time, but they had to part with their daughter.

"They chained the black princess to a white limestone cliff by the sea. The waves broke with roar against the high cliff, and the pearly foam tenderly licked the feet of the doomed sacrifice.

"Suddently the wide sea foamed, whirlpools appeared in it, and out of the deep rose the terrible Whale. Its jaws opened greedily, a flame sparkled from its small, fierce eyes, and gray smoke curled up from its ears. Its scaly tail curled along the water in rings, fiercely cracking along the waves.

"Now the monster noticed its new sacrifice, and its eyes flamed still brighter. Closer and closer it swam, cleaving through the sea.

"Meanwhile, among the white heaps of clouds the brave hero Perseus was flying on winged sandals. He had recently with his magic sword cut off the head of the terrible Medusa, and the winged horse Pegasus had arisen out of her blood. The glance of Medusa turned into stone anybody who dared to look her in the eyes. But Perseus had fooled Medusa and fought with her, looking not at Medusa herself, but at her reflection on his polished shining shield.

"And now Perseus flew happily. He carried Medusa's head with him, and loathsome snakes squirmed on it instead of hair.

"Suddenly Perseus looked down below and saw, on the seashore, a beautiful maiden chained to the white cliff and a terrible monster was moving towards her.

"Perseus instantly rushed to the fight and directed the glance of Medusa's head at the Whale. The Whale turned to stone and became a rocky island washed by the blue sea. Meanwhile, Perseus freed Andromeda and led her to the palace where the king happily presented Andromeda to him as his wife. And then the appeased gods placed the images of all the participants at this event in the sky."

You will find in the sky near one another the constellations Cepheus, Cassiopeia, Andromeda, the Whale, and Perseus with Pegasus. And one of the stars in the constellation Perseus was called Medusa's Head for a long time... .

In ancient times, constellations were understood as arrangements of bright stars characterized by the figure they make up if mentally connected by straight lines.

However, we almost never see in these characteristic figures any resemblance to those objects by whose names the ancient astronomers called these figures.

If one speaks about constellations as about "figures", then Ursa Major and Ursa Minor resemble each other, but because of their long tails they do not resemble a bear; they rather remind a dipper or pan.

Nowadays, the word constellation means a whole territory in the *sky* within definite boundaries. All the stars seen in this area of the sky are referred to that constellation. However, in *space* if for simplicity the borders of the constellations are assumed to be a circle, the constellation includes all the stars that lie within a cone whose apex starts at our eyes and the lines forming the cone extend to the borders of the constellation. Some of the stars in a given constellation in space are farther from their neighbours in the constellation than from stars that appear to us to be on the opposite side of the sky. Finding the constellations, however, helps us to be oriented in the sky.

Brightness and Names of Stars

The writer Ivan Popov in his novel *Toward Dawn* wrote:

"You see, Paul, the star... there, right above the chimney... the bright one... . Do you know what it is called?"

"I don't know."

"Let it be my star and yours, our star... the star looked at me... . Let's call it Pitatseya. You don't know what Pitatseya is? I don't know either... it seems I made it up myself, although, perhaps, I heard it somewhere... Pitatseya... ."

But a thousand years ago people had already given names to the brightest stars. The Romans, Greeks, and Arabs gave them names. The names of bright stars remain with us today, Sirius, Aldebaran, Vega, Antares, and others. Weaker stars are known by the letters of the Greek alphabet approximately in the order of their luminocity, within each constellation separately. So Sirius is α of the Big Dog, Vega is α of the Lyra, Castor is α of the Twins, and so forth. Still weaker stars are only designated by a number in the catalogue that registers them by their celestial coordinates and index of brightness.

Two thousand years ago the greatest astronomical observer of antiquity, Hipparchus (fl. between 146 and 127 B.C.), sorted the stars by brightness into six groups, into six star magnitudes. The brightest (about 20 all in all) he called the stars of the first magnitude, weaker ones were stars of the second magnitude, and those hardly visible to

the naked eye stars of the sixth magnitude. The greater the brightness of a star, the lower is its stellar magnitude.

Hipparchus's proposal turned out to be convenient, and it has been basically preserved and refined. Fractional magnitudes have been introduced for stars of intermediate brilliance, for example 2.15 or 3.47. They are denoted $2^m.15$, or $3^m.47$. When a star is taken as a model of a first-magnitude star, then stars $2\,^1/_2$ times brighter are called stars of zero magnitude, and stars $2\,^1/_2$ times brighter still, stars of minus one magnitude (because the brightness of Hipparchus's brightest stars did not turn out to be strictly the same).

In comparison with first-magnitude stars, sixth-magnitude stars are 100 times weaker. Stars of the seventh and eighth magnitude are correspondingly $2\,^1/_2$ and 6 times weaker. The very weakest stars, which can only be seen clearly in a photograph obtained with the most powerful modern telescope, are 2,000 million times weaker than first-magnitude stars, and their apparent magnitude is as low as 24.

Addresses of Stars in the Sky

When you go for a stroll through an unfamiliar city, a plan of the city will come in very handy and if you look for an acquaintance in the city, you will need his address. When you go for a stroll across the starry skies, a star chart (a plan of the skies) will be of use to you and to search for a star in the sky, you will need its address in the sky. The address of a star is its coordinates in the sky. We are still speaking about the *apparent* place of a star in the sky, not about its position in space. To an observer the sky appears to be a ball or sphere at which he looks from the inside and at the centre of which he is located. Half the sphere is hidden below the horizon and all the heavenly bodies appear to be equally distant from us, i. e. as though they were located on the surface of this sphere. Actually, this is of course not so, but by thinking of all stars as being on the surface of the celestial sphere (with an indefinite radius), we can conveniently fix the apparent positions (coordinates) of the stars. Knowing these positions we can easily find these stars in the sky. After finding a star in the sky, we can then note its location on the star map.

If we, being at the centre of the imaginary celestial sphere, mentally extend through our eye a plane parallel to the plane of the Earth's equator, then it will intersect the sphere in a large circle called the celestial equator.

Imagine a line passing through our eyes and parallel to the axis of rotation of the Earth; it will intersect the celestial sphere at points called the poles of the Universe. The

northern pole happens to be close to a bright star, called the Pole star. As is well known, it is easily found from a group of seven bright stars arranged in the form of a dipper and called the constellation Great Bear.

Both on Earth and in the celestial sphere, it is possible to determine the position of any point with two coordinates. On Earth these coordinates are the geographical latitude and longitude. Latitude is reckoned in degrees starting at zero from the Earth's equator and increasing along the arc of a meridian either towards the North or towards the South pole. The latitude of a point is the number of degrees from the equator to it. Longitude is measured along the equator from a meridian accepted as the base meridian to the meridian passing through the given point.

Instead of geographic latitude for the celestial sphere *declination* is used. Declination is measured in degrees, like latitude, from the celestial equator towards one of the poles of the Universe.

Instead of geographic longitude *right ascension* is used on the celestial sphere. It is the angle between some base celestial meridian and the meridian passing through the point in the sky.

The declination and right ascension of a body are its coordinates, its address, in the sky. The address directory of the stars is a catalogue containing their celestial coordinates. There are so many inhabitants in the starry heavens that address directory of the stars frequently gives only an address, without the star's name. In that case the star is simply known by its number in the catalogue.

The overwhelming majority of "sky-dwellers" are settled, i.e. they "live" at the same address all the time, but there are some wandering bodies, the celestial nomads. These are the planets and comets. "Planet" is the Greek for "wanderer". It became clear that the wandering bodies move in the solar system, i.e. relatively close to Earth. Their apparent location in the sky changes as a result of their own movement in space and also as a result of the movement of us, observers, with the Earth. The remaining bodies also appear to move, but their movements are hardly noticeable because they are so far away from us. Thus if you are working in a garden you will notice the slightest changes in the position of people working in adjoining gardens, but a remote gardener will appear to be staying in the same place all the time.

Distances between bodies on the celestial sphere, like distances between points on any sphere, are measured in degrees. These angular distances should not be confused with the distances between the same bodies in *space*. In the sky two stars may appear to be close to each other, but in space they may be many times farther apart. The apparent, angular

diameter of bodies is the angle at which their true (i.e. linear) diameter is seen from Earth. For example, the angular diameters of the Moon and Sun are almost equal, subtending about $^1/_2°$ when seen from the Earth. The Moon is small but close to us; while the Sun is huge but far from us. It is useful to remember that a segment of the circumference approximately equal to 1/57 of the radius appears from the centre to subtend an angle of 1°, and a segment of 1/206,265 of the radius appears to subtend an angle of 1″. This permits the linear diameter of a body to be calculated when its angular diameter and the distance to it are known.

For example, if from a distance of 150,000,000 kilometres the Sun has an apparent diameter of $^1/_2°$, then its linear diameter is 150,000,000/114, i.e. about 1 $^1/_2$ million, kilometres.

Eyes, Ears and Hands of Astronomers

It sometimes happens that a sceptic, upon learning about the latest discovery of astronomers, shakes his head unbelievingly and says: "But who saw it, who measured it, and who touched it?!" For this reason, I will begin by telling you how astronomers see, how they measure, and how they "touch" the heavenly bodies, i. e. what kind of "eyes", "ears" and "hands" they have.

The eyes of astronomers – brown or blue, gay or meditative – are the usual eyes of people. But in our times astronomers cannot contribute much to science unless they take advantage of the glass eyes of a telescope.

However, when people come to an observatory and wish to look through the telescope or at least to see how astronomers look in the telescope, they are often in for a disillusion. They are told that the astronomers here do not actually examine anything through telescopes and that the observatory does not have a single telescope through which an astronomer would look with his eyes alone.

Nowadays, a photographic plate usually replaces the human eye, and it would, perhaps, be more correct to call *the camera* the astronomer's eye. But not only it alone receives the light of remote stars. There are many other accurate and sensitive devices, which can determine the brilliance and colour of heavenly bodies better than eyes and can feel the heat released by a heavenly body better than hands. They receive from the stars the radio waves and other radiations that are invisible to the naked eye. These "instrumental sensations" could be called the "sixth sense" of astronomers.

Let us go through a large astrophysical observatory designed primarily for the study of the physical nature of stars and find out what can be seen there.

General Survey of an Observatory

On a dark winter night the sky sparkles with little lights, and their twinkling invites us to study the stars. But the astronomer showing us around frowns in dissatisfaction. "Again strong shimmering," he says. "That means that the air is not still, and therefore the images of the stars in the telescope will vary and shake. It will be difficult to examine them." It is dark in the observatory, and in the past astronomers of all countries have sometimes heard guests, perhaps familiar with photography, inquire sympathetically: "I suppose you probably, have to use a flash to photograph the stars?" But to illuminate with a flash of magnesium even the nearly Moon would be somewhat difficult; when photographing stars using their own light alone, it is necessary, on the contrary, carefully to guard the photoplate from outside light. Therefore, sometimes even someone smoking in the vicinity must be careful!

The astronomer will probably hasten to anticipate another question some guests have: "We do not predict weather, that is the job of meteorologists." Glancing round the observatory complex, you notice that besides the one or two large buildings there are large and small towers with domes, which can be turned at

Inside the Byurakan Astrophysical Observatory near Yerevan, USSR.

will, spread over the territory. The domes have open apertures-hatches through which air freely passes and telescopes look at the sky. A clock mechanism slowly turns the telescope, following the daily rotation of the sky. When photographing with the big telescope, an astronomer only occasionally looks in a small telescope, which is fastened parallel to the large one, in order to check whether the clockwork is operating correctly and whether it is necessary to adjust the position of the telescope or to change the speed of the mechanism.

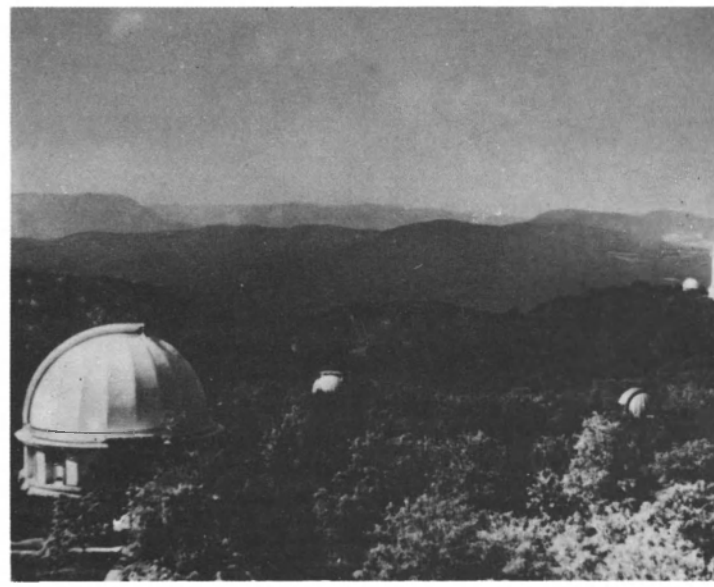
It is cold in the tower. But is it really impossible to observe through a telescope in the warm? Alas!... The glass from which telescope objective lenses are made must be of the best quality as far as its homogeneity is concerned. Its surface must have the assigned form and cannot deviate from it by more than one ten thousandth of a millimeter. Any ordinary glass in front of the objective, even the best brand, would so ruin the images given by the telescope that it would be impossible to make out anything. The observer likewise cannot be dressed in a suit with an artificial heater because streams of warm air would come out of the suit, and in passing up through the hatch they would further spoil the images of the stars, which are fairly spoiled as it is by disturbances in the atmosphere, that permanent "enemy" of the astronomers.

The presence of this "enemy" is the reason why they can show you the Moon in the large telescope only with the same magnification as in a small telescope, i.e. 150-200 times, rarely 300 times, and never over 500-600 times. Although a telescope with a 5-metre diameter mirror can magnify up to 20,000 times, such a magnification cannot be used because of the interference caused by the disturbances in the Earth's atmosphere.

"Why then build larger and larger telescopes?" you may ask. An astronomer would answer:

"The larger the diameter of a telescope the more light it collects, the more stars are visible in it, the farther stars which can be observed with it, and the smaller the time lost on photographing them."

It would never enter anybody's head that senior astronomers look through the largest telescope and junior astronomers use the smaller ones. Even apart from the fact that in most cases now people do not actually look through the telescope and instead photograph through it or attach another sort of device to it,

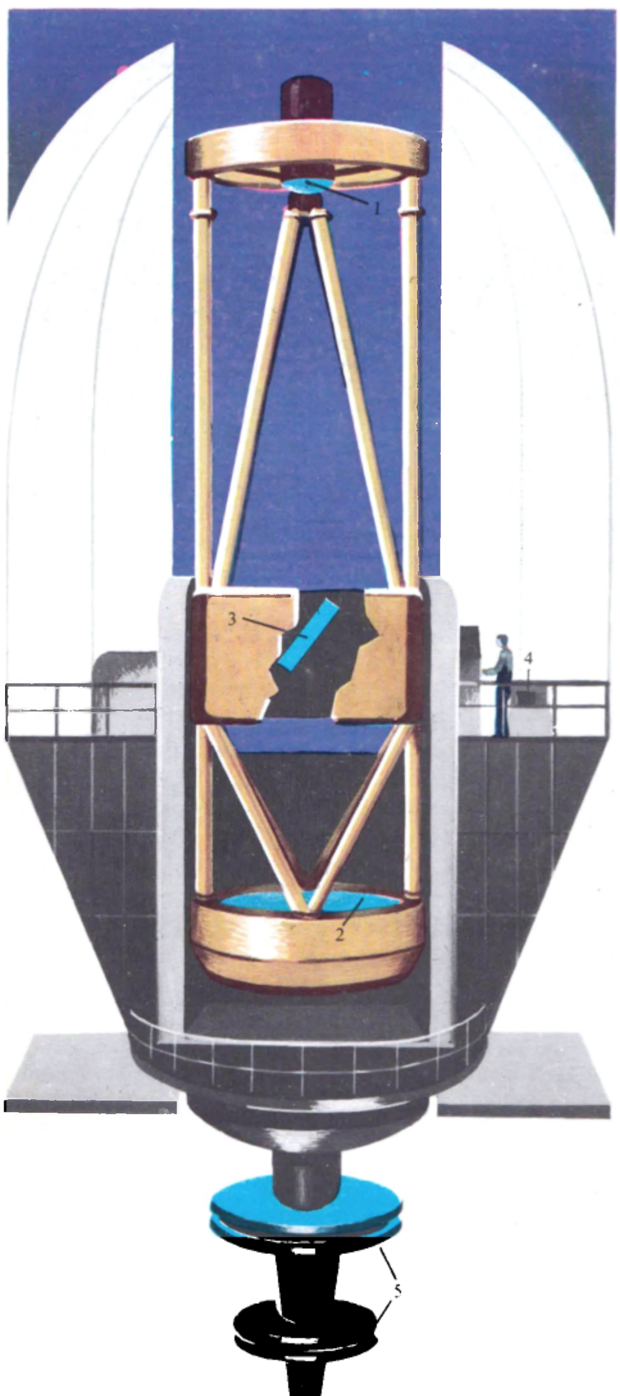


The Crimean Astrophysical Observatory, USSR.

telescopes are assigned not according to the rank of astronomers but according to the jobs they perform. Telescopes have very different characteristics, and their quality and value are not determined by size alone. If he had to accomplish a certain scientific goal, an astronomer would not exchange a special small telescope for the largest that would be less suitable.

The reason is that some telescopes, those which give large images, only permit a small sector of the sky to be photographed at one time, and they require a long exposure to obtain a "portrait" of a comet or another heavenly body. But other telescopes can photograph large sectors of the sky, even though on a small scale, and they can give a small but bright image of a comet with a tail. They also permit an astronomer to use short exposures. Telescopes which are specially adapted for photography are called *astrographs* (astro + graph) although the name *photographer* (photo + graph) is given not to devices but to people. This sometimes leads to confusion.

Photographs of the sky, particularly glass negatives, are cherished at observatories like gold. Over time an entire "glass library" is formed from the negatives, preserving the history of the sky. Each negative is an important document. By comparing old photographs



of the sky with new ones, we can discover changes in the sky, like variations of the luminosity of stars, their mutual displacement, and so forth.

Besides the usual sort of telescope you might see some outlandish enormous structures, trusses, metal nets and hundreds of poles on frames. You are told that these are radiotelescopes, but we will discuss them later.

Optical Telescopes–Astronomers' Eyes

The main job of a telescope, we repeat, is to collect as much light as possible from the heavens. Astronomers have never had enough light from stars to analyze them in every way. That is why they are always seeking to increase the diameter of their telescopes. At the same time they are also trying to increase the scale of the images they can obtain of the bodies. The scale depends on the length of the telescope or, more accurately, on its focal length, or the distance from the objective to the place where the image of a body is received. The larger the image, the better it can be examined.

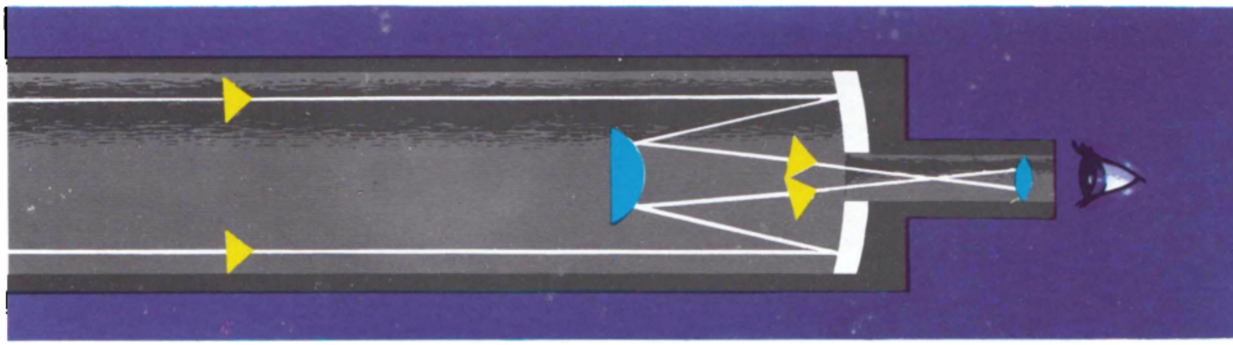
As the size of a telescope increases, all the difficulties encountered in its manufacture grow because its weight increases proportionately to the cube of its diameter, but the accuracy of manufacture remains the same. This accuracy is determined by the fact that the glass of the objective must be everywhere equally homogeneous and that it must be ground and polished to within 1/10,000 of a millimetre. In addition, the mount must be so perfect from a mechanical point of view that the telescope can move, following the daily revolution of the sky, without permitting deviations greater than 2-3 hundredths of millimetre.

Just imagine a bulky and cumbersome structure, which can be compared in weight to a locomotive, and which is made with such accuracy and moving smoothly at the observer's will!

Telescopes rotate, following the daily revolution of the sky, around an axis pointed at the pole of the Universe, being guided by a driving clock.

For two centuries there was a fight between two types of telescopes: *refractors* (with a convex refracting glass objective) and *reflectors* (with a concave re-

The 6-metre reflector telescope of Zelenchuk, USSR, to observe stars down to the 24th apparent magnitude:
1–auxiliary mirror; *2*–primary mirror; *3*–diagonal mirror; *4*–spectrograph; *5*–drive.



The Cassegrain reflector telescope. A small image formed by the concave mirror is reflected from the convex mirror and observed through a hole in the concave mirror.

flecting mirror). At the end of the last century the reflectors won out for large instruments because opticians began to make mirrors out of glass instead of metal, as before. The glass was then covered with an extremely thin layer of silver or aluminium to reflect light. A mirror does not need the fine and strictly predetermined sort of glass as is required for an objective, through which light passes all the way. Also it is only necessary to grind one surface, not four as in an objective, which usually consists of two glasses. Reflectors are not only cheaper and easier to make, but they can be made in dimensions that turned out to be unattainable for refractors.

In the 20th century they have not been able to make objectives larger than those made at the end of the last century (one metre). They have not even been able to make objectives of the same size. Meanwhile, reflectors of this magnitude are made without particular difficulties, and after World War II in California (USA) a giant telescope with a mirror 5 metre in diameter was built.

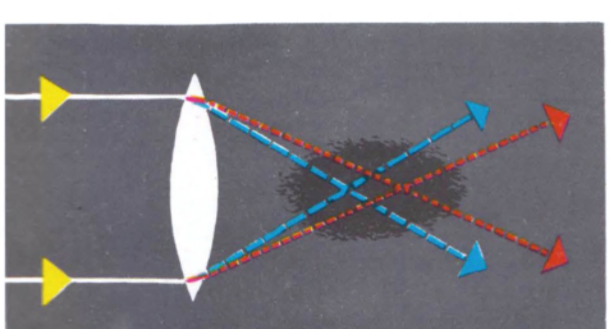
In Russia Jacob Bryus (1670-1735) made the first reflectors at the beginning of the 18th century. Since then a number of large observatories were built in Russia, especially after the Great October Socialist Revolution.

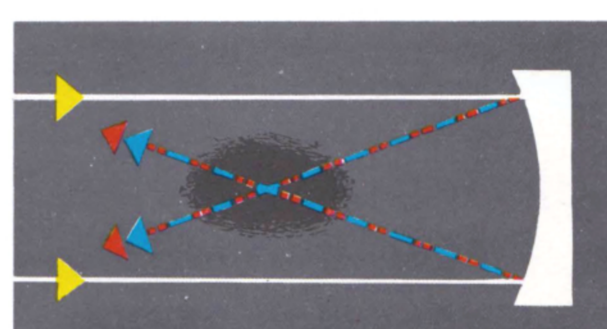
These "fortresses" of Soviet astronomy suffered greatly from the invasion of the fascists. Pulkov Observatory, the "astronomical capital of the world" (as it was frequently called abroad following the American astronomers Gould and Newcomb) was destroyed during the blockade of Leningrad.

Thanks to the concern of the government, the observatories ruined by the fascists were restored and in fact reequipped better than they had been before. Several new observatories were also built.

The Soviet Union now has its own optico-mechanical industry. This industry did not exist in tzarist Russia but was built up during successive five-year plans

The different optical paths (the lens and the mirror) for light of different wavelengths.





10^{-9}
10^{-8}
10^{-7}
Visible
10^{-3}
10^{-2}
10^{-1}
1
10
10^{2}
10^{3}
10^{4} cm
X-rays
Radio
Gamma-rays
Ultraviolet
Infrared
Ionosphere
Stratosphere
Troposphere

and after the war became particularly strong. However, simply enlarging the types of telescopes that have long been known or simply expanding the telescopes retains all their deficiencies.

A reflector collects all the rays of different colours into a single focus to give an uncoloured image, but the trouble is that it is poorly suited for photography. Just a little way from the point of the sky at which the telescope is aimed the images in focus turn from points into something resembling birds; they smear out in such a fashion that you "do not want to look at them".

An astronomer though does want, and often even has, to photograph a large sector of the sky, e.g. a whole constellation, at once. What is to be done? Won't an ordinary photographic objective help? We still cannot make very large high-quality objectives (from several lenses) to photograph stars. Moreover, the glass in them would absorb too much light as a result of their thickness. Simpler objectives do not give sharp images of a large sector of the sky, although they are better than reflectors. In addition, every objective always gives a slightly coloured image because it does not collect all the different coloured rays coming from a star into a single point.

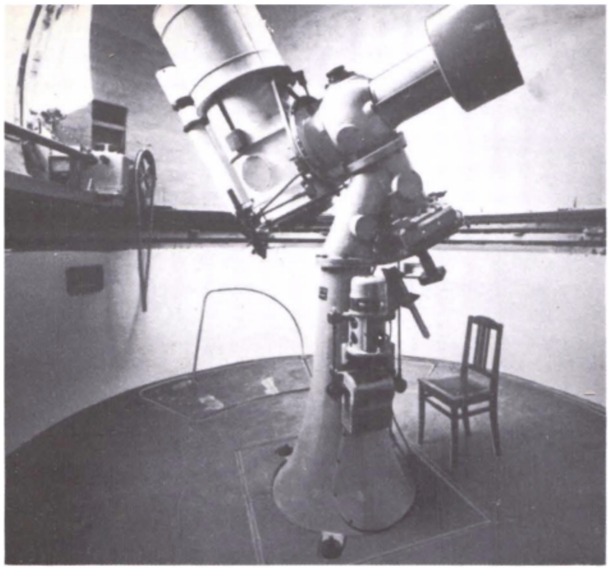
Now the simple reflector that has the above-described defect is not concave like the inside of a ball (spherical), it has a somewhat different shape called a paraboloid. A parabolic mirror is considerably more difficult to make than a spherical one, and even though many other shapes have been tried there has been little progress.

When they began to ship out the starving, but still steadfast, Leningraders from the besieged city, a very tall man lay on a berth in one of the trains and thought. Many interesting and useful thoughts had been born in this passenger's mind, but now he was thinking about the fate of his offspring. This offspring was the proposed mass production of school telescopes, which until then had not been produced in the USSR. It was proposed to make a mirror (small reflector) telescope. But, in addition to the war, which had stopped the production of telescopes, this reflector promised many difficulties. Its aluminium mirror would tarnish from the action of the air and dust, and the schools would reject them... so mused the passenger. "It would be well to preserve the mirror from this danger by protecting it from the front by a plane-parallel glass," thought Dmitry Maksutov, for the passenger was indeed this outstanding Soviet optician.

"But then," Maksutov continued to reflect, "it is possible to fasten the small mirror in the reflectors to the first piece of glass. This mirror will throw the rays backwards or to one side where the rays will be focused and where the eyepiece will be placed. This will eliminate the requirement for special holders for the small mirror, which would spoil the image. But in one of the small telescopes the small mirror is convex or concave. Why then not replace the plane-parallel glass by a meniscus, i. e. a convex-concave glass, in such a fashion that its central silvered part is itself a small mirror with the required curvature. "Of course," thought Maksutov, "it is necessary to take a meniscus with surface curvatures which, like a plane-parallel

The Pulkovo Observatory near Leningrad, USSR.

Until fairly recently astronomy has only been concerned with optical investigation of celestial bodies using the light that comes to us through the atmosphere. Besides this optical "window", there is a radio range in which radiation from space comes to us. The strides made by radio engineering and radiophysics have enabled astronomers to take advantage of this range, and so after World War II radio astronomy emerged. The barrier that the Earth's atmosphere presents for short-wave and infrared radiation can be penetrated using balloons (on a limited scale), rockets, satellites and spaceships with instruments on board. In recent years the whole electromagnetic spectrum has been put to use. From this ever growing body of evidence astronomers have been extracting valuable, even surprising results.



A Maksutov telescope at Pulkovo, USSR.

glass would not distort the reflections given by a non-spherical mirror otherwise... ." And right here Maksutov made his discovery. It is possible to take an easily manufactured spherical mirror that distorts the image, so to speak, in one way, and to make a meniscus in such a fashion that it distorts the image equally in the opposite direction. As a result, the errors in the system cancel out, and the image would be faultless in form. Owing to the small thickness of the glass meniscus there will be no noticeable differentiation in the position of the focus for different rays. Thus Maksutov recounted the course his thoughts followed.

A meniscus telescope can be used instead of a usual reflector, and it will also be better and several times shorter, i. e. more convenient in use. A meniscus telescope can also be made into a camera of huge dimensions for *photographing* large sectors of the sky. For a number of years, small meniscus telescopes of the Maksutov type have been produced for schools. Also, some large meniscus telescopes have been produced for scientific goals. The largest of these, with a diameter of 70 cm, was established in the mountains of Georgia, in the Abastumani Observatory.

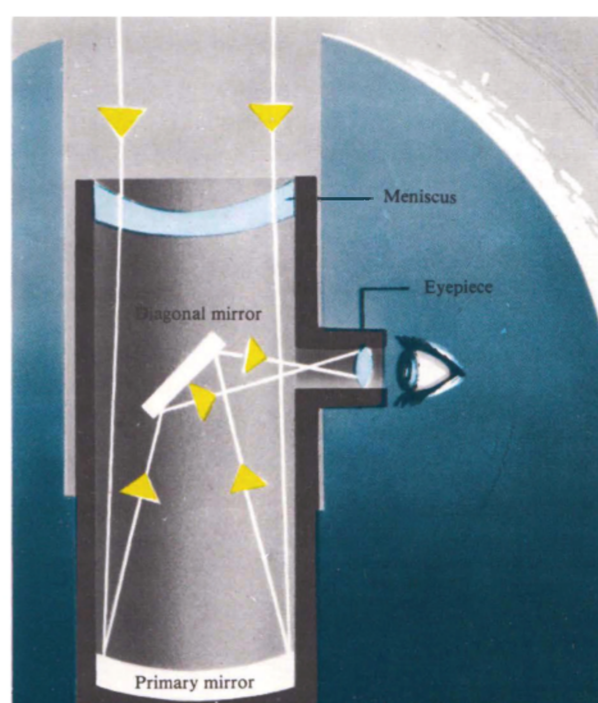
The strides made by the Soviet optico-mechanical industry enabled it in 1967 to assemble the world's largest telescope with a 6-m parabolic mirror.

A small telescope can be manufactured at home. Any accurate and patient person, including a schoolboy, can make a real telescope with a mirror 10-15 cm in diameter, without much trouble or expense. This is incomparably easier than you might think. Unfortunately, it is considerably more difficult to make a good mount for such a telescope, because this takes various materials and also abilities to work with metals and wood, talents few lovers of the sky's secrets can boast.

Before World War II there was no suitable reflector at Moscow University, so I decided to manufacture it myself. My dear readers, I was as unskilled in grinding 26-cm mirrors and making telescope mounts as most of astronomers. So I had to become an amateur optician and amateur designer, mechanic, and turner. The job was a success, although it would not take the cake for elegance. Later the University of Moscow got more powerful telescopes, so I presented my telescope to Saratov University. There it was additionally equipped with a clockwork and now it is put to service at the satellite observation station. It is also used for scientific work and demonstrations of planets and stars to students and the general public. Many amateurs have made their own telescopes of even larger diameters and better workmanship. So, you see, as the Russian proverb says "it doesn't take a god to burn a pot".

The mirror is made by moving one thick glass disk

The Maksutov telescope.



round and round and from the centre to the sides over another, similar one with some wet abrasive dust between the pieces of glass.

Telescope's Auxiliaries

Precise measurements with the aid of auxiliary instruments give us fascinating and detailed information about the physical nature of stars and planets, which simply examining and photographing them cannot.

There are *measuring devices* to measure distances on a plate with an accuracy of $^1/_{1000}$ of a millimetre. These are used in a laboratory at an observatory to measure the positions and dimensions of heavenly bodies.

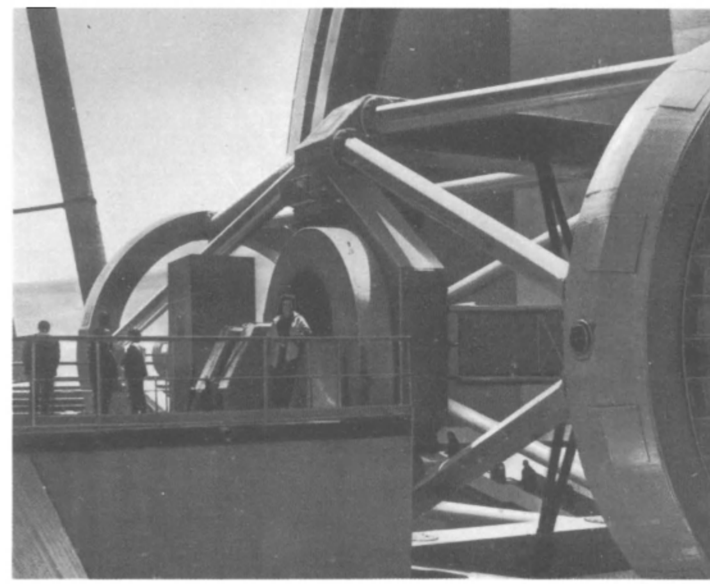
Although the devices are "auxiliary" to the telescope, the data obtained by them are often more valuable than those we obtain by observing directly through it.

Photometers, including the photoelectric type, are devices to measure the brilliance of a body. They are based on the same principle of turning light into electric energy that is used in a movie, for instance.

Various types of *microphotometers* are used to measure the darkening on photographic negatives, making it possible to measure the brightness of the stars.

As he is accompanying you away from an observatory, an astronomer might tell you, "Please do not light a match until you get over 300 kilometres away from the observatory because today we are going to measure the heat received from the stars, and the apparatus would react to your match." If he should say this, it would only be half in jest. We all know that the stars warm us very little, particularly on a frosty night. To learn how to measure the similar insignificant heating of the Earth by the stars is the greatest achievement of modern instrument-making. In fact, the *thermoelement* that measures the heat from the stars has an exceptional sensitivity; but, of course, every possible measure is taken to prevent the slightest spurious radiation from reaching the instrument. Therefore, you could not only light a match harmlessly within 300 kilometres, you could even light it without leaving the tower of the telescope.

A thermoelement consists of a junction of two wires of different metals. If the junction is covered by carbon black and heated, then an electric current will flow in the wires. Of course, the weld is not covered by



An altazimuth reflector with a 6-metre mirror in a tower at the Zelenchuk Observatory, USSR.

carbon black to make it dirty, but so that all the energy reaching it will be absorbed and turned into heat. In order to be convinced of the ability of a black, dull surface – like that of carbon black – just change your white hat in summer for a black one.

In order to measure the heat from stars, two 0.01-mm diameter wires are used. Their mass is 0.03 mg and the electric current arising in them is measured to within 3×10^{-11} amperes*. Only an accurate apparatus like this can measure the heat reaching us from the bright stars. The most heat comes to us from the bright red star Betelgeuse in the constellation Orion – 7.7×10^{-11} calories per square centimetre per minute. Even if we were to collect this heat for a year with the aid of a 2.5-m concave mirror, we would only gather enough energy to heat up a thimbleful of water by two degrees. Just try to calculate how long it would take to bring a tea kettle to the boil in this fashion.

Astronomers have other ways to verify the

* 10^{11} designates a number that is a 1 followed by 11 zeroes, and 10^{-11} means 1 divided by 10^{11}. The "order" of any number is its approximate size, and so 3×10^{11} is said to be of the order of 10^{11}, and 3×10^{-11} is of the order of 10^{-11}.

thermoelement evidence. Incidentally, thermoelements have actually proven to be excellent means of measuring small amounts of thermal energy under terrestrial conditions.

Astronomy uses cinema techniques a little bit, e. g. to study fast changes on the Sun. But in most cases the light of celestial bodies is too faint to allow filming, whereas the variations in them are too slow to warrant filming anyway. The movements you see in astronomical movies are shot using dummies – globes and balls – showing the planets and stars. Television began to be used in astronomy fairly recently. So one cold clear night in Moscow the Moon's image obtained through a large telescope was directly transmitted to the general TV audience from the observatory of Moscow University. Television has also been used for scientific purposes. N. F. Kuprevich at Pulkovo used television to study the Moon's infrared (thermal) radiation, and astronomers at the Crimean Observatory saw stars on the screen of a television system coupled to a telescope that were fainter than those that could be observed or photographed through the same telescope. But TV techniques in astronomy are as yet complicated and not reliable enough. Television was used extensively after the advent of automatic interplanetary stations. So we observed numerous "close-ups" of the Moon and Mars and even magnificent panoramas of the lunar surface.

Late in the 1950s new techniques appeared permitting the radiation from space to be studied in formerly unaccessible spectral ranges. These methods, or branches of astronomy, began to be called astronomy with various attributes, just as earlier still the term "radio astronomy" was coined. Now they speak about "infrared", "X-ray", "balloon", "satellite", or "neutrino" astronomy. The expansion of studies to include the infrared, far ultraviolet, and X-ray ranges is extremely important for a better understanding of the nature of the stars and the processes occurring there. Up until fairly recently the shielding of these parts of the spectrum by the blanket of our atmosphere was a major obstacle. Today photomultipliers, electrooptical transducers, improved photoplates, and radiation detectors can be sent up using remote-controlled balloons and high-altitude rockets. Placing equipment above the atmosphere has also allowed the energetic primary cosmic rays to be studied.

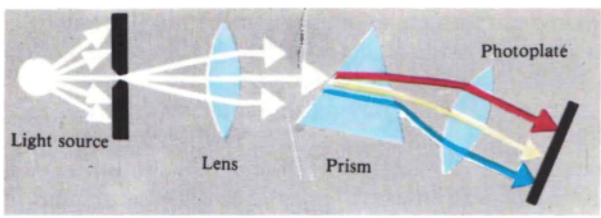
Experiments are under way to "catch" neutrinos, the elementary particles produced by nuclear reactions on the Sun and other stars. They are evidence of processes deep within the stars about which so far we only have theoretical predictions.

The spectrum analysis of the light coming from heavenly bodies brings us the most data of all. The spectra of stars characterize their physical and chemical nature and many other qualities. Just what a spectrum is and what it looks like cannot be described briefly and so the next section is devoted to the subject. Until you have thoroughly mastered the principles of spectral analysis, you cannot understand most of the conclusions of astrophysics, that is the physics of heavenly bodies.

Spectroscopy and Spectra

Rays of light are the emissaries of their master, the source of the light. When broken down into a spectrum, they tell us about the physical condition of the body which sent out these rays. It may be as far away from us as one likes, if we can just obtain sufficient light from it to photograph its spectrum and study it. Such a photograph of a spectrum is called a *spectrogram*, and is obtained with the aid of a *spectrograph*.

The construction of a spectrograph is pictured in the accompanying figure. The main part is a three-sided prism made of a transparent substance, usually glass, that refracts the light of different wavelengths (of different colours) differently: the shorter the wavelength, the more the refraction. Thus green rays are refracted and bend farther toward the base of the prism than red rays, and violet rays bend even more than green. Light is waves, and, depending on the wavelength, we receive the impression of one colour or another. The length of a wave is expressed in ten-millionths of a millimetre, which are called ångströms. So the wavelength of green colour is about 5,000 ångströms (or about five ten-thousandths of a millimetre). The slightest change in the wavelength corresponds to a change in colour (although infinitesimally small and unnoticeable to the eye). Some light sources emit light of only one wavelength; others emit light consisting of rays of several wavelengths, of which some may be bright, others weak, and still others may be intense but still invisible to the eye. For example, infrared or heat rays, which have a very large wavelength, are invisible to the



The principle of the prism spectrographs.

naked eye. Also invisible are ultraviolet rays, which possess a wavelength shorter than 4,000 ångströms, and X-rays, which are incomparably shorter waves, but like ultraviolet rays, also act on photographic plates.

The X-rays and "long" ultraviolet rays of celestial bodies do not reach us because they are absorbed in our atmosphere by the layer of ozone (a gas whose molecules consist of three atoms of oxygen).

Radiation with wavelengths longer than infrared rays are used for radio communication. All the waves of the spectrum, from radio waves to X-rays, are called electromagnetic waves.

Usual spectral analysis deals with the spectrum from the infrared to the ultraviolet radiation. Only the middle of this range is visible to the eye. Yellow and green light affects human eyes the most. Therefore, the brightness of one section of the spectrum or another, as it appears to the eye, still does not characterize the *energy* of the radiation at a given wavelength. Brightness for the eye depends both on the energy contained in a given part of the spectrum and on the sensitivity to it. The same applies to photoplates.

If a narrow metal band covered by carbon black is moved through the spectrum, it will absorb all the energy falling upon it and turn the energy into heat. Depending on how much it is heated, the conductivity of the metal changes; therefore, by measuring the current flowing through this changing electrical resistance, we can determine the true distribution of energy over the entire spectrum. This instrument is called a *bolometer*.

The colour of a source is dependent on the wavelength and the intensity of the light it produces. For example, hot sodium vapour radiates almost exclusively at the wavelength of a bright yellow and therefore it is yellow. Most of the energy of hot mercury vapour is radiated at wavelengths corresponding to green and dark-violet.

Some sources, for example the filament of an electric lamp. radiate light at all possible wavelengths (without exception or interruption); and so its spectrum is called continuous. The colour of these light sources depends on the distribution of energy over different wavelengths, i. e. on the distribution over the spectrum. For example, if the most energy is released at the red end, then the colour of the light source is red, but if in the ultraviolet, which is invisible to the eye, then the colour of the source is determined by the brightest place in the *visible* part of the continuous spectrum or the combination of colours in the brightest places. For example, it is well known that white is created by a mixture of all the colours taken in a particular proportion or in some proportion of two colours, for example yellow and blue, or red and green. This mixing of light of different *colours* must be distinguished from the blending of *paints* in painting.

But let us return to the figure. In order to break light down into its component parts, it is necessary to direct a pencil of parallel rays at the prism at a precise angle (depending on the characteristics of the prism). This is done by a *collimator*, which is a tube with an objective facing the prism and having a short slit at the other end. The slit is parallel to the edge of the prism and is located where the light from the prism and falling on the object lens in a parallel pencil would be focused into a point. This point is known as the main focus of the objective. There is a principle in optics which states that the rays sent out by such a lighted slit will leave the objective and fall on the prism in almost a parallel pencil.

At the slit of a spectrograph the telescope's objective gives almost a point image of a star or the extended image of another celestial body, like a planet. In the spectrograph itself the refracting prism disperses light to yield a rainbow. The shorter the wavelength, the greater is the refraction. These rays enter the camera at different angles. The camera's objective produces an image of the slit on the photoplate in one place or another, depending on the angle of incidence of the light, which in turn depends on the wavelength. Therefore, we obtain on the photoplate a small band consisting of a number of lines parallel to each other. These are images of the slit, and each image is formed by the rays (from the source) of a definite wavelength.

This coloured band, consisting of a number of lines, is the spectrum. From the place occupied by a line in the spectrum, we can determine the wavelength of the component colour. For brevity, scientists speak about the determination of wavelengths of the lines themselves in the spectrum. If a spectrum contains all wavelengths, without interruption, then the countless merging images of the slit will give a *continuous* spectrum. It contains all the colours of the rainbow (or rather a rainbow contains all the colours of the spectrum). Consequently, light sources that emit only certain wavelengths yield a spectrum in the form of a number of separate lines. This is a *line* spectrum. When the spectral lines are bright, this is an *emission* spectrum. But sometimes in front of the source of continuous spectrum there is a substance that absorbs light in separate, definite wavelengths, thus removing the energies of those wavelengths from the continuous spectrum to produce places in the spectrum devoid of light, i. e. dark lines. This is an absorption spectrum, which is also a line spectrum.

Spectral Language

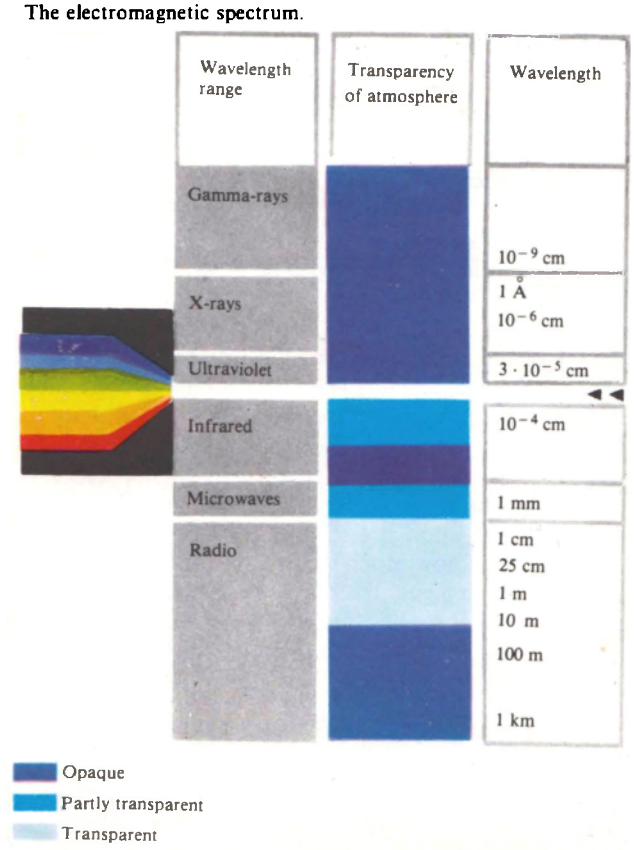
We shall now see how the certificates of the heavenly bodies (their spectra) are read, we shall study the language of spectra.

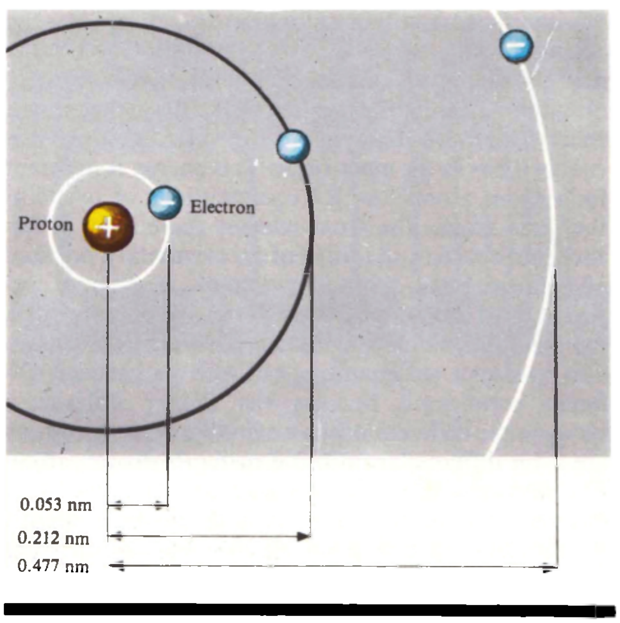
Even the general appearance of a spectrum tells us something about the luminescence of a source. Experiments tell us that a continuous spectrum is given by either hot solid or hot fluid substances or gases in which there are many free electrons, those smallest particles of electricity. A small layer of an extremely hot and dense gas or an extremely thick layer of a thinner gas can give a continuous spectrum.

Thus, on the one hand, continuous spectra are produced by the filament of an incandescent lamp or by molten steel and, on the other hand, by the gases forming the surface layers of the Sun and the stars. In the laboratory it has been possible to turn small wires into gases containing multitudes of electrons (by burning the wires with strong electric currents), and just when they vaporize the gases emit continuous spectra.

An atom is the very smallest part of a chemical element, i. e. a portion that cannot be further divided chemically into smaller components. Nevertheless, an atom is an intricate thing. It constitutes a complex system of particles. The atoms that make up different chemical elements contain different numbers of particles and are constructed differently, but it is impossible to break them up or recompose them using chemical means. Once only nature could transform them, but now it is possible to change them in laboratory by methods of physics. You can imagine an atom as consisting of a nucleus with electrons orbiting around it. The nucleus has an inherent weight and an inherent positive electric charge which is equal to the total negative electric charge of the atom's electrons. The nucleus's charge determines the chemical properties of the atom. If it collides with another atom, electron, or the smallest particle of light – a quantun or photon – our atom may lose one of its electrons. It then acquires a single positive charge that is not balanced off by the positive charge of the electron that has just left its master. The atom then becomes an ion, or an ionized

The electromagnetic spectrum.





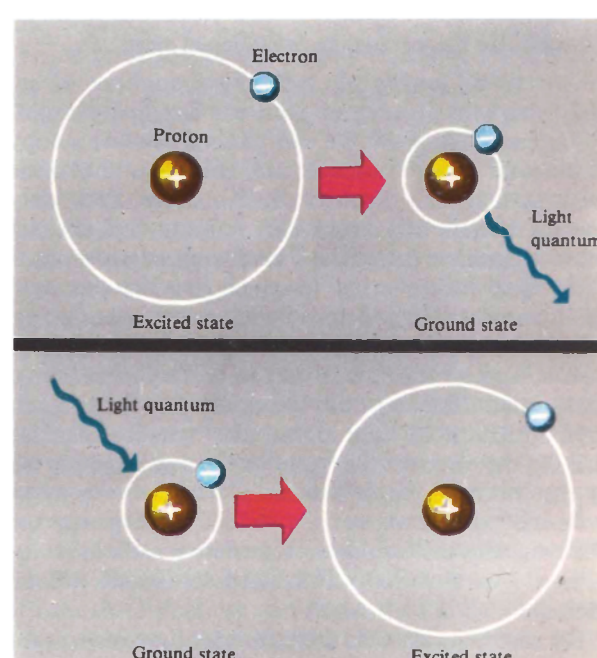
The first permissible orbits of the hydrogen atom.

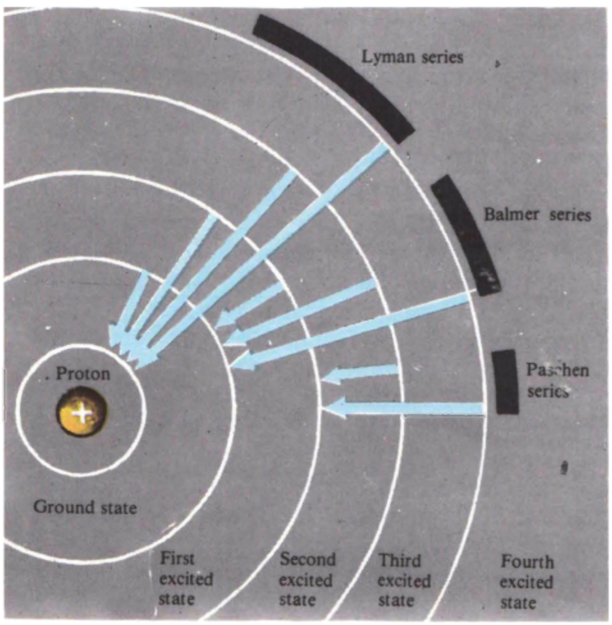
atom. If a second electron is also removed from the atom, then it becomes a doubly ionized atom with a double positive charge.

Atoms do not feel "angry" about these and similar injuries and losses in collisions; however, they constantly cherish the hope of restoring their disturbed "household" and refilling it by capturing free electrons. Electrons liberated from their "serfdom" are called free, but they are always threatened with being captured again because their negative charge is attracted by the positive charge of ions. Woe to the slow-flying free electrons! They are easy game. Meanwhile, fast electrons hurtle around ions without danger, and it should be clear that the sparser the atomic population per unit of volume (the thinner the gas), the easier it is for them to slip away from captivity, preserving their freedom and thus maintaining the high ionization of the gas. But the more particles crowd in one place, the more frequently they collide and reunite.

It is known from physics that the higher the temperature of a gas, the quicker its particles move, the more energetically and frequently they collide, and the greater the number of atoms that ionize. The velocity of the particles determines how often collisions can take place. Based on this, theory can predict what the distribution of energy over a continuous spectrum will be for a given temperature. The radiation of an *absolutely black body*, so called because it is capable of absorbing all the energy falling upon it, has been studied theoretically best of all. An absolutely black body is not only the best absorber of radiation, it also has the greatest emissive power at a given temperature. The radiation of the filament of an electric lamp or that of the interior of a hot stove greatly resembles the radiation of a black body. It is possible to create an artificial body that will resemble an absolutely black body even more closely. On heating it you would be convinced, as physicists are, that the energy distribution in its spectrum depends on the temperature, in a way corresponding to theory. It turns out that the Sun and stars possess almost the same properties of emission and absorption as a black body. It is therefore possible to determine their temperatures from the energy distribution over their spectrum, a subject to be discussed later. Thus, do not be confused that the Sun is taken to be a black body, and an absolutely black body at that! Black coal still behaves like a black

Light emission and absorption by the hydrogen atom.





Lyman, Balmer and Paschen series of hydrogen, named after the physicists who contributed the most to the theory of the hydrogen atom. The Lyman series belongs to the ultraviolet range and is due to transitions of electrons from excited states to the ground state. The Balmer series is due to transition to the first excited state; and the Paschen series, to the second excited state. Only the Balmer series lies in the optical range.

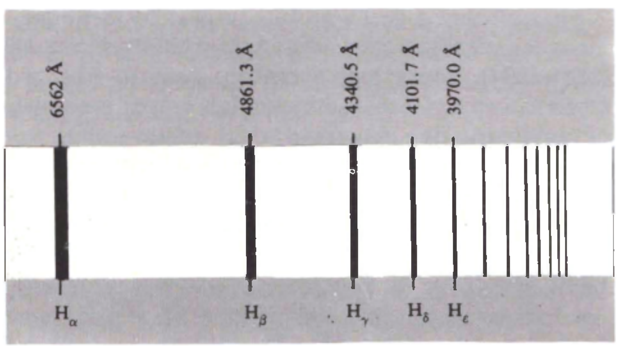
body even when it is burning hot and dazzlingly bright.

But back to atoms. We can picture them as a copy of the solar system in miniature. However, unlike the solar system, the electrons in an atom may only occupy definite orbits and may only move from one orbit to another in jumps. The energy of an atom is determined by the orbits in which its electrons stay, the innermost orbits corresponding to the least energy.

In order to shift an electron in an atom to an orbit with a bigger radius, it is necessary to expend some energy, and this energy can be imparted to the electron by a quantum of light or another moving particle striking the electron. The greater the energy imparted by the collision, the farther the electron moves away from the nucleus. At certain, rather large energy the electron will be torn away from the nucleus completely, i. e. it will fly away and the atom will be ionized.

However, the electron does not stay long in an orbit far from the nucleus. An inexorable law of nature requires that a hundred-millionth of a second later the electron must jump back again to an orbit closer to the nucleus, giving off part of its energy in the form of radiation. This energy is equal to the difference in the energy of electrons staying in the outer orbit and the energy it has in the inner orbit. This energy is different for various atoms and for combinations of orbits in the same atom. The atom releases the energy difference into space in the form of an elementary portion, of light quantum, whose wavelength, or colour, corresponds to its energy. So a hydrogen atom releases a red quantum, a blue quantum, etc. The cadmium atom also releases a red quantum, but with a somewhat different wavelength because the energy differences between the different orbits attainable for its electrons are a bit different from those in the hydrogen atom. The electron shells in ionized atoms are arranged in a different manner from unionized (neutral) atoms, therefore the quanta given off by them are different from those of neutral atoms. Still, for clarity we can picture the situation as though each jump of an electron to an orbit closer to the nucleus produces a short sound of a strictly definite tone. The vast number of atoms in a hydrogen gas (even not a very dense gas) radiate all the different quanta attainable for hydrogen, i. e. the spectrum radiated by hydrogen consists of a number of wavelengths and spectral lines inherent in hydrogen. The same applies to the other chemical elements and their combinations. This makes it possible to determine the chemical composition of a source gas from its spectrum. This has been known a long time experimentally, but it was not until recently that it was fully explained, when the theory of atomic structure was developed.

An atom is like a piano that can only produce certain sounds, certain notes, which also, incidentally, correspond to definite lengths of sound waves. Each type of atoms has its set of "notes" (the spectral lines emitted by it). Unlike a tuner, a spectroscopist cannot change the pitch of these atomic "notes", but he can notice the difference between two almost identical notes better than a tuner, and he can determine from them the type of atom (piano) producing them. The spectra of ionized atoms are different from the spectra of neutral atoms. If all the electrons were removed from a nucleus, then the atom would lose its capability to radiate because it would not have any electrons left to jump from orbit to orbit, and its energy would not



The Balmer series.

change. It would, as it were, change from a piano into a simple wooden thing without strings.

Thus, qualitative spectral analysis is done from the wavelengths of the lines in spectra. It is used with equal success for heavenly bodies and in many fields of practice on earth such as in physics, chemistry, geology, biology, medicine, and metallurgy.

Thin gas releases bright spectral lines characteristic of it, and issues its own set of light notes, a melody of its own, as a result of the excitation of its atoms by collisions with other particles of matter or with photons. Just as chords may go together in a melody, so the so-called multiplets, or series of lines which usually appear together, go together in a bright line spectrum of an atom. For example, the pair of yellow lines in the sodium spectrum and an entire series of lines in the hydrogen spectrum represent such chords. They always appear together although, for various reasons, other groups of lines in the spectrum of the same atom may be absent. However, the same gas acts differently if a hotter source of continuous spectrum is placed behind it.

Being cooler, the gas will absorb the energy of the continuous spectrum that falls on it from outside. Its atoms are only able to absorb certain of spectral wavelengths. Therefore, only those wavelengths from the continuous spectrum that the atoms would in other conditions issue themselves will be weakened by being absorbed. Upon absorption, the gas causes dark lines to appear in the continuous spectrum that correspond exactly to the lines in the emission spectrum of the gas. The absorption spectrum and the emission spectrum of an atom are like the negative and positive of a photograph.

Of course, as we already know, after absorbing some energy and becoming excited to a state with greater energy, the atom must instantly return its aquisition. However, the light quantum may be radiated by the atom in any direction whatsoever. The vast number of atoms, having received energy from one direction (from where the source of the continuous spectrum is located) discharge the energy in every direction. As a result, only a part of the energy contained in the continuous spectrum will come to an observer looking at the source of the continuous spectrum through the thin gas. The energy at the appropriate wavelengths will arrive in an even more weakened state, and the observer will see a dark line in the spectrum. The picture we have described is called the diffusion of light by atoms.

It is as though the quanta radiated in a continuous spectrum are a shower of coins that a crowd standing far away from you throws to you in a game. Now imagine that a dozen boys come between you and the crowd and begin mischievously catch only the nickels and quarters leaving the other types. After catching the coins, they throw them in any direction. It is clear that among the coins that reach you there will be fewer nickels and quarters although some will reach you anyway.

The more the absorbing atoms there are in the path of the continuous spectrum, the greater the energy will be absorbed and the darker, or as they say more intense, will be a dark line of the spectrum.

In our example, the more boys who "break the rules", the greater the loss you will suffer in terms of nickels and quarters. Knowing the absorbing capability of atoms (i. e. how well the boys catch the coins), we can calculate the number of absorbing atoms (number of boys) there are in the path of a light ray (shower of coins). Thus, a quantitative chemical analysis can be carried out as well on the basis of the intensity of the lines in the absorption spectrum.

In astronomy the sources of continuous spectrum are hot surfaces of stars and the Sun, which consist of tremendous masses of ionized gases. The surfaces are surrounded by hot, though cooler, gaseous atmospheres. The scattering of light in these atmospheres causes dark lines in the spectra of stars and the Sun. From these lines we can make qualitative and (from their intensity) quantitative chemical analysis of the atmospheres of the stars and Sun.

If we recall that stellar atmospheres consist of many chemical elements and many types of atoms, and each type gives its own series of lines (as if it were producing its own melody consisting of different chords), then it should be clear what a good "musician" the spectroscopist must be in order to distinguish meaning from the cacophony of spectral lines (the basic medley of notes belonging to various arpeggios and chords). It sometimes happens that when a star's spectral lines are being identified some notes are out of tune. Then it is necessary to struggle to find out to which melody and to which piano (atom) the false note really belongs.

It remains for us to tell you how spectra reveal the movement of celestial bodies. Recall Doppler's principle, which should be familiar to you from your school physics. It states that if a source of oscillations moves with respect to us, then the wavelength of these vibrations, as perceived by us, changes. As the source comes closer, the wavelength shortens, and as it moves away, the wavelength becomes longer. In the case of sound vibrations, a speeding locomotive's whistle gives a common example of this. As the train rushes towards us, the sound of its whistle is higher, and the pitch lowers sharply when the train, after spewing fumes on us, begins to move away quickly.

In the case of light oscillations, the wavelengths in the spectrum change. However, even at speeds of hundreds of kilometres per second, it is impossible to detect changes of colour in the spectrum because the change in the length of the waves is too small. Only in science fiction stories can a car driver state that a red light of a traffic signal appeared to him to be green because he was travelling so fast towards it. For this to be the case he would have had to be moving at over 60,000 kilometres per second.

Only shifts of the lines of spectra – changes in their wavelengths, $\Delta\lambda$ – can be detected. According to the Doppler effect the velocity v of a source with respect to us is

$$v = c\frac{\Delta\lambda}{\lambda}$$

where λ is the normal wavelength of spectral lines and c is the velocity of light, which is equal to 300,000 km/sec. As a light source draws nearer the observer, the spectral lines shift towards the violet end of the spectrum. As it moves away, they move towards the red end.

That all this is really so was proved by the prominent Russian astrophysicist Aristarkh Byelopol'sky (1854-1934), about half a century ago. In Pulkovo Observatory he set up in his own laboratory a number of rapidly rotating mirrors in which a light source was reflected so that its image moved at a velocity approximating the velocities of heavenly bodies at which it is only possible to notice the Doppler shift of lines in a spectrum with certainty. This cleared up any doubts about the truth of Doppler's principle.

Astronomers' "Ears" – Radio Telescopes

Any heated body sends out electromagnetic waves – ultraviolet, visible, thermal, and radio – although in different proportions, depending on the body's temperature and properties. We have already said that for the so-called absolutely black body – the ideal radiator – Planck has derived a formula relating the energy in its complete spectrum to its temperature.

The constituent particles of a heated body all participate in thermal chaotic motion. The kinetic energy, i. e. the energy of the motion of the particles, is transformed into the energy of the electric and magnetic fields in collisions. The electromagnetic energy output grows as the "swiftness" of motion that determines temperature increases. The Sun, stars and clouds of thin interstellar gas all produce thermal radiation, which we can measure. But there are also other causes of radiation.

An electrically charged particle changing velocity produces an alternating electromagnetic field, i. e. the particle radiates energy. The velocity of an electron is changed when a positive charge attracts the electron as it flies past it. The power of the radiation produced is minute, but the gases in space contain such enormous numbers of electrons and protons that their combined radiation yield is significant. Also, electrons and protons change their velocity in magnetic fields, which makes them spiral along its lines of force, thus sending out electromagnetic radiation, particularly in the radio range. This radiation is named cyclotron radiation after the device used in nuclear research – the cyclotron – where this radiation was first observed. Cyclotron radiation has also been found in space.

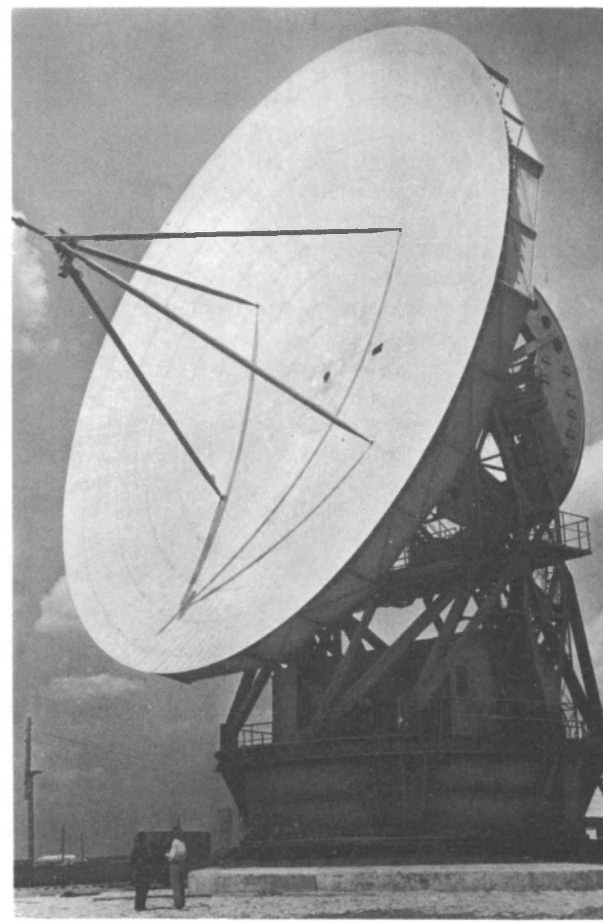
All the above kinds of radiation form a continuous spectrum in the radio range, just like the one observed in spectral analysis. Unfortunately, the continuous

spectrum of the radio radiation from space does not come to us without interruption due to absorption by the Earth's atmosphere. Or rather, radio waves are absorbed by the higher, electrified layers of the atmosphere, known as the ionosphere.

A "window" in the ionosphere only passes wavelengths from 16-20 metres to 1 $^{1}/_{4}$ centimetres. Even the 1-mm waves are poorly transmitted, although they are not stopped by the ionosphere, but by atmospheric water vapour. These wavelengths are neighbours of thermal waves, which are known to be absorbed by water. It is only through this radio window that we peer out into, or if you wish, "hear" the developments in the radio range beyond the terrestrial atmosphere. What is more, it is only in this range that we can send out our radio signals from the Earth into space. Space probes travelling above the Earth's blanket can receive and transmit radio signals at any frequency, but now spacecraft cannot yet have on board the powerful radio equipment and electric power supplies required to study the very faint and distant cosmic radio emissions.

Radio broadcasters must fight noise that interferes with their programmes, say a concert, especially if the broadcast is coming from far away or their station is not powerful enough. Radio astronomers, too, must deal with noise, which comes from a variety of processes occurring in space. We know that the space between celestial objects is not empty, it contains ionized particles and magnetic fields. Noise is also produced by our atmosphere and the radio set itself. So the major task of the radio astronomer is to distinguish the radio frequency of some object against the background of noise.

Nowadays, radio-frequency signals are recorded automatically. The pen of the recorder traces the "level" of the incoming signal on a paper tape. Noise appears as a serrated streak, whilst the signal as a peak over it, the higher the peak, the stronger the signal. Processing these signals is a problem and the specific conditions of the recording session should be taken into account. Not infrequently the device records the signals coming not only from the source of interest but also from other spurious sources. Astronomers thus listen to signal and noise. The antenna of a radio telescope is, as it were, the "ear" of the astronomer. The larger the "ear", the more energy arriving to us from distant worlds it catches. The antennas of radio telescopes come in a wide variety of designs. So there is a radio telescope that resembles an optical reflector in that it has, as its key element, a mirror, only it is made of metal. This is a huge dish concentrating the radiation at its focus where the feed, a small antenna, is located. The energy from the feed comes along a waveguide to a room housing the receiving apparatus. From this room the dish is trained at the required area of the sky by just pressing a button actuating the electric drive of the telescope. The radio telescope differs from its optical cousin in being absolutely unhampered by clouds.



The radio telescope at the P. N. Lebedev Physical Institute of the USSR Academy of Sciences.

We know from optics that for a mirror to collect radiation exactly at its focus, the irregularities of its surface must be less than the wavelength of the radiation received. With light the wavelength is under one micrometre, and with radio waves it is of the order of centimetres and metres. The mirror of a radio telescope therefore permits markedly larger surface tolerances. Unlike optical telescopes, large radio telescopes are much easier to fabricate. The size of radio telescopes is important for two reasons.

First, the larger telescopes collect more energy. Second, they have larger resolutions than smaller ones, i. e. they are able to distinguish two individual sources at small angular separations. But the resolution drops with wavelength. However, now radio astronomy can sometimes achieve accuracies larger than with the best optical observations. A large solid mirror is a very massive thing. A rule of thumb is that if a mirror has defects that are about 10 times smaller than the wavelength, they are of no consequence.

Accordingly, if a radio telescope's dish is designed for 1-metre waves it can have holes up to 10 centimetres in diameter. This allows it to be made in the form of a metal net, instead of a solid structure which reduces the dish mass and costs, and makes it easier to manufacture.

The radio telescope at Byurakan, USSR.

So radio reflectors in recent years have reached diameters of almost 100 metres, their mass amounting to hundreds of tonnes even with lattice-type structures. Clearly, with such enormous sizes radio telescopes are placed in the open, not in a tower. The larger and heavier a telescope, the more difficult it is to turn it with the required accuracy and to follow the daily revolution of the skies. Therefore, it pays sometimes to limit the mobility of a telescope fixing it so that it only observes the sky near a meridian or even only near a zenith.

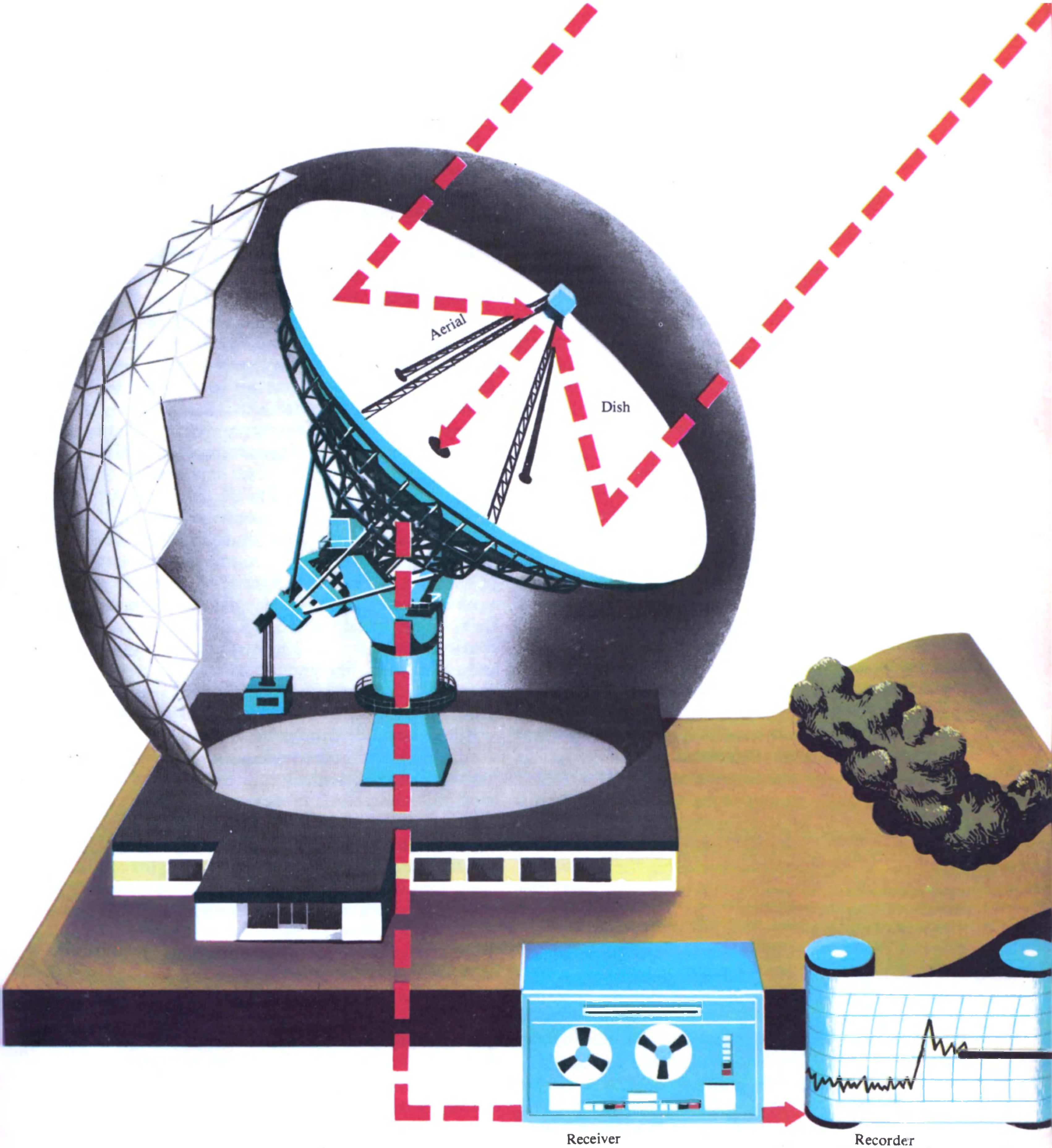
Which is the best and largest telescope today? I cannot answer this question.

First, new, ever larger telescopes are being constructed, and at the moment of writing the situation may have already changed.

Second, radio telescopes vary widely. They have different possibilities of motion, some are large, but designed to catch, say, only very long radio waves, the others are more universal. The largest reflector is absolutely fixed and mounted at the bottom of an extinct volcano in Puerto Rico. Its diameter is as large as 300 metres! It is directed at the desired object by shifting a cabin suspended on ropes between masts 135 metres high.

There are radio telescopes in the form of sets of matched small reflectors. At the Pulkovo Observatory the radio telescope has an arc-shaped construction whose ends are separated by 120 metres. Another Soviet radio telescope is a system of reflectors arranged along an arc 600 metres in diameter. Also, there are telescopes (interferometers) in the form of a cross with sides hundreds of metres long, and antenna telescopes in the form of a plane frame with numerous dipoles fastened on it. If you happen to see these monstrous structures, you would hardly believe that they are radio telescopes. The interferometric method provides the greatest precision. The object is observed simultaneously by two large radio telescopes separated by... the diameter of the globe. A larger separation, i. e. larger base, is as yet impossible. This technique enables angles about $0''.001$(!) to be measured, which is impossible with optical telescopes.

Thus, radio engineering, too, connects the Earth with the starry heavens. Speaking in the words of the



The principle of the radio telescope.

Soviet poet A. Kovalenkov:

The quiet of the frosty night,
The forest has cast a spell on the air,
The ground is buried in crystal,
The stars converse with it.

The space of fathomless fire-flies,
Heavenly bodies of eternal fairy tales,
So high, so far,
Above the branches of the birches and the elms... .

The Earth speaks to them about us,
And, beneath the tense antenna,
We stand, not dropping the eye,
At a confrontation with the Universe.

Astronomers' "Fingers": Radar and Laser or... Is It Possible to Feel Planets and Illuminate the Moon?

Another important method of promise is radar (abbreviated from Radio Detecting and Ranging). It can spot an object capable of reflecting radio waves we beam to it. Radar investigations have made great strides during and after the war as a military tool but then it found an application in one of the most peaceful branches of science, astronomy.

Using suitable equipment, e. g. a concave reflector, we now can collect radio waves into a parallel beam. The energy in the beam is scattered insignificantly and can reach a remote object and upon reflection return back to the radar station to be received. The radio waves are generally sent as short powerful pulses. When the echoes return we determine the direction of the object in the skies, and measuring the amount of time required for the radio waves to travel to and from the object. Our high-accuracy apparatus works out the distance to the object, as it is known that electromagnetic waves travel at the constant speed of 300,000 kilometres per second.

The first astronomic spin-off of radar techniques was in lunar studies. True, the Moon is clearly seen and its distance and apparent position in the sky are known from numerous observations and calculations. But it was extremely interesting to check the earlier results by radar methods.

L. I. Mandel'shtam and N. D. Papaleksi, two Soviet physicists, showed in 1928 that radio studies of the Moon were possible (you see that here we also need a theory). These experiments required a powerful radio transmitter because of the significant energy losses at great distances involved, and so it could only be carried out in the USA and Hungary in 1946.

Higher powers and improved sensitivity of receiving apparatus made it possible in a short period of time to "touch" more distant celestial bodies with radio beams. As early as 1964 the first radar studies of Venus, Mercury, Mars, and finally of more distant planets Jupiter and Saturn, were carried out in the USSR, and also in Great Britain and the USA.

Of especial significance were the radar studies of Venus to determine its distance and the semimajor axis of the planet with higher precision than using astronomical techniques. This also improved our knowledge of our major "yardstick", the Sun-Earth distance. It is called the astronomical unit (AU) and is used to measure distances in the solar system. The earlier method consisted in measuring as accurately as possible the distance to the nearest celestial object and working out from it the solar distance using the fact that semimajor axes of the orbits of all the solar satellites are related by the Kepler third law to their periods, which can be determined very accurately from observations. The currently accepted value of the astronomical unit is 149,597,870 kilometres.

But radar offered ways other than just measuring the distances within the solar system. With sufficiently narrow beams of incident radio waves and sufficiently accurate timing of the echoes we can measure the distances to various points on a planet's surface. So radar appeared to be a valuable tool for exploring the surface features of the planet such as the heights of mountains, topographic peculiarities, etc. This is especially important when the planet is shielded by clouds.

The nature of radar reflection is dependent on the roughness of its surface. If the surface is smooth, the radar returns will only come from the central portion of the planet's hemisphere facing us. The echo will show just a peak. The whole of the exposed hemisphere will only contribute to the reflected signal when the incident wave strikes the slopes normal to the direction of its incidence. This will "smear out" the reflected signal in time, thus enabling the average roughness of a planetary surface to be estimated. In addition, radar allows us to establish the rotation of a planet about its axis.

If a planet rotates (and its axis does not face the observer), then one "edge" of the planet approaches us, while the other recedes. According to Doppler's principle the wavelength of the signal coming from these edges will change: in the first case it becomes shorter, and in the second longer, the wavelength coming from the centre remaining unchanged. So the "width" of the echo, as an interval of wavelengths or frequencies, will be larger than with the initial signal. Signal broadening will be the larger, the faster the linear speed of a planet's rotation. So we can work out the linear speed, and hence the period, of a planet, knowing its diameter. It is this technique that has allowed us to determine the utterly slow rotation of Mercury and Venus. The earlier methods relying on fixing some spot on their surfaces and on spectra appeared unreliable.

Lastly, radar is an indispensible tool for measuring the heights and speeds of "shooting stars", but we will be looking at this subject later in the book.

Is it possible to illuminate the Moon? It is no use to illuminate it at full moon – it is bright as it is. But can we illuminate a "new" moon, if only a tiny spot? Isn't it too far? It appears that we can now illuminate our satellite, although only a small piece of it, and, of course, not with a search light but with a laser. The laser (abbreviated from Light Amplification by Stimulated Emission of Radiation) accumulates light and yields a thin light pencil of strictly one frequency (called monochromatic light). Lasers already have extensive applications in science and technology. It became possible to send laser beams to some places on the dark side of the Moon and illuminate them sufficiently to be seen through telescopes.

In 1970 the self-propelled moon-car *Lunokhod 1* carried a French-built laser reflector consisting of a series of silver-plated quartz lenses. The device reflected laser signals, which were then received by the observatories in France and the Crimea. The laser studies of the Moon, due to its precision, make it possible to study the drift of terrestrial continents, the polar motion, and a number of questions of space geodesy and celestial mechanics. Light signals beamed by a laser installed on a lunar roving vehicle crawling over the dark side of the Moon will be visible from the Earth (a usual search light would be invisible).

The Hunt for High-Energy Cosmic Rays

The terrestrial atmosphere, fortunately, protects the Earth from lethal radiation emanating from the Sun and other sources elsewhere in the Universe. The shorter the wavelength, the larger the energy of quanta of electromagnetic radiation. Take, for example, X-rays, which are known to have formidable penetrating power. But even they lose their energy in endless collisions with atoms and molecules during their long journey through the Earth's atmosphere and, at least in their initial form, they do not get as far as the Earth's surface. Meanwhile, they are of immense interest, being associated with atomic processes, and so the investigation into their emission helps astronomers understand the processes in their parent bodies. Yet shorter wavelengths are produced in radioactive decay, this radiation being termed gamma-rays, or γ-rays.

In 1962 an X-ray counter installed on a high-altitude rocket found a source of X-ray radiation in the constellation Scorpion. From the end of 1970 to the beginning of 1975 about 200 X-ray sources in space were found. X-ray astronomy is now making great strides. Specifically, our Sun is also studied in X-rays.

Thus, space messages are now being received over the entire range of electromagnetic radiation, from optical radiation through thermal and, further, to radio waves, and on the other side through ultraviolet to X-rays.

Another kind of energetic rays are "cosmic rays". These are now understood in a narrower sense to be "corpuscular rays" coming from space. These are elementary particles, mostly protons and electrons, ejected by some powerful processes with speeds approximating the speed of light. Therefore, their kinetic energy and penetrating power are beyond believe. Long exposures to them are known to be a danger for astronauts even near the boundaries of the terrestrial atmosphere. Just like X-rays, cosmic rays beyond the Earth's atmosphere are now studied by satellite-borne apparatus.

Yet another way of research relies on a more traditional technique, balloon astronomy. Sometimes fairly large telescopes designed mostly for solar observations are raised into the stratosphere with the help of strings of balloons.

Through Soviet Observatories

In the past Russian astronomy, which was justifiably proud of its outstanding representatives, was poorly equipped, if we exclude the first-class Pulkovo Observatory.

Besides Pulkovo, through the vast territory of the Russian Empire there were only a few small university observatories with rarely more than 3 astronomers, working in each. The tzarist government did not favour science with its attention.

Socialist rule has radically changed the situation and the country is now covered by a whole network of strongholds of astronomy. The outposts of these strongholds, which were moved far to the south by the oldest veterans of astronomical science in Leningrad and Moscow, were quickly turned into the well-equipped observatories of the sister republics. In the tzarist Russia's former colonies and semi-colonies, there was often not even a single full-time astronomer, but now there is a local cadre of fully-fledged members of the Soviet astronomical community.

The old Pulkovo Observatory is now adequately fitted with modern Soviet-made instruments. It has Maksutov and Slyusarev telescopes, a large horizontal telescope for solar studies, a new Sukharev meridian circle, spectral and photoelectric equipment, powerful radio telescopes, and so forth.

The Sternberg Astronomical Institute of Moscow University moved in 1954 its main part to new spacious premises in the Lenin Hills, not far from the remarkable tall building of Moscow University. In the Crimea under the dark southern sky in 1958 the southern base of Moscow University was commissioned. New, larger telescopes (up to 125 centimetres in diameter) have been installed there.

The Crimean Observatory of the USSR Academy of Sciences at Simeiz, which had been destroyed during World War II, was restored, and a considerably larger new observatory was built near Bakhchisaray. They are both now united into the Crimean Astrophysical Observatory. One of them commands a magnificent view of the sea and cliffs of the southern slopes of the Yayla Mountain Range, and the other faces the mountain chains of the northern slopes of the same mountain massif.

A new 2.6-m reflector was set up there, as well as a cine camera for photographing the Sun, powerful cameras and devices for photographing the night sky, radio telescopes, and many other instruments for "dissecting" the atmospheres of heavenly bodies.

In the foothills of the majestic Aragats mountain in Byurakan the observatory of the Academy of Sciences of the Armenian SSR arose above the valley. Well equipped and located in southern latitudes high up in the mountains, it is one of the best Soviet observatories, boasting 1 and 2.6-m reflectors. Its specialty is stellar astronomy.

Also in the mountains, above the famous resort of Abastumani in Georgia, an astrophysical observatory is functioning that is mostly involved in the study of the colour of remote objects and the study of interstellar dust. It has the largest Maksutov telescope, with a diameter of 70 centimetres.

At Shemakha in Azerbaijan a new observatory with a 2-metre telescope has been opened in 1967.

On a high ridge in the Northern Caucasus there has been built the Special Astrophysical Observatory with the world's largest (6-metre) reflector and down in the valley a huge radio telescope RATAN-600.

A large mountain observatory near Alma-Ata in Kazakhstan was the first to demonstrate the efficiency of Maksutov telescopes by successfully using a 50-cm Maksutov telescope to study the fine details of interstellar gas clouds.

University and other observatories were expanded or constructed, which was made possible due to the young, but rapidly developing optical industry of the Soviet Union.

"Smart" Satellites

Man has now learned how to make satellites, ones that differ from the satellites "manufactured" by nature. Natural planets are hurtling along in space, deaf, dumb and blind.

But man-made satellites, more like asteroids in size and motion, are "smart". They "see" and "speak". They also "hear" as they receive the radiation coming from celestial bodies and radio messages sent to them from Earth.

After World War II the USA and USSR began to launch high-altitude probes with automatic apparatus. When such rockets rise up to the thinnest strata of the atmosphere, they not only leave all the clouds behind, but also the layer of ozone gas, which absorbs the far-

ultraviolet radiation of stars thus forbidding its study. Even these short-lived rocket flights gave us much valuable information derived by spectrographs and other equipment. In addition to data on the structure of the atmosphere itself they carried out the first studies on the far-ultraviolet and X-ray spectrum of the Sun, thus increasing our understanding of it.

On October 4, 1957, the first artificial earth satellite in the history of mankind was launched in the USSR. This was a triumph of science and technology because for the first time man had overcome the pull of the Earth's gravity and created an artificial moon; it was not just a swiftly falling, high-altitude rocket, but a small laboratory that circled the Earth for several months.

Konstantin Tsiolkovsky (1857-1935), a pioneer in rocket and space research, would not dream of the striking development of aeronautics and astronautics. Hundreds of satellites are now plowing through terrestrial space, automatic probes have been launched to Venus, Mercury, Mars, and Jupiter. They have returned a great deal of evidence concerning these planets that has resulted in a revision of our concepts of the planets. Some of the interplanetary stations softland modules, or place in orbit round the Moon or the planets artificial satellites for long-term studies. The information gleaned is returned on command from the Earth. We will look at the evidence derived in this way in the later chapters.

You Yourself Can Observe and Study the Universe

Although scientific research and making discoveries generally requires extensive special training, sophisticated expensive apparatus, and a great deal of reading in a number of languages, this by no means suggests that amateur astronomers have no chance of making their contributions. All the more is it accessible to anybody to observe celestial events and based on what one has read in books to find interesting developments in the sky, to analyze their progress, so increasing one's understanding of them.

Even merely watching the starry abyss that opens up before you on a clear night will give you a profound satisfaction. On such a night a person, as it were, sees and feels infinity and eternity. Just seeing the starry skies alone is a deep aesthetical experience.

But, apart from just contemplating the sky, the amateur can study astronomy, and the stars at first hand and even make his own contribution to the progress of the science. This assures of boundless cognizability of nature, and gives satisfaction to the mind. You should only choose problems that are equal to your skill, knowledge, and capabilities.

I answer very many letters from my readers. It is a pleasure to answer an inquisitive reader who wants to know more or who has provided a competent description of some phenomenon he is having difficulty in understanding. But there are, unfortunately, cases where the correspondent states outright that, say, the law of gravitation is all wrong as it stands, and that only the one he has invented (exactly, *invented*) is the true one. Strangely, such proud inventors forget that the law of gravitation has been tested for centuries in innumerable contexts and by innumerable people. For example, it is common knowledge that spaceships are launched and reach their destination in full accordance with the law of universal gravitation. Just because the authors of popular-science books devoted to complicated physical problems try to give a simplified version of some theory, some people think that they too can create a new science as simply.

You can see and trace many interesting things through a telescope, binoculars, and just by unaided eye. You can make your own telescope and use it for observing and photographing. You can take pictures of the Moon, Sun, eclipses, comets, planetary wanderings, and constellations; you can also make your photographic map of the sky.

In recent years many amateur astronomers have produced beautiful pictures of comets, which have enormous scientific value. We even see the beginnings of amateur radio astronomy, although in a small way. Some amateurs who are conversant with electronics even succeeded in sending a radio signal to the Moon and receiving an echo.

Their observations of the so-called noctilucent clouds, and visual, photographic and radio observations of "shooting stars" may also be of scientific value. With an average-size telescope, even a home-made one, you can make useful drawings of the changes in the planets. You can observe variations in the brightness of a variable star or discover a new star, which is generally done precisely by amateurs. For example, an amateur discovered in August of 1975 a

new star in the constellation Cygnus, the Swan. These are just the things for which professional astronomers have no time.

How this is to be done, and what programme must be followed we cannot describe it here, of course. We refer the interested reader to numerous manuals for amateur astronomers. Try to "read" the starry skies, the great book of nature, and try to grasp the meaning of what you have read.

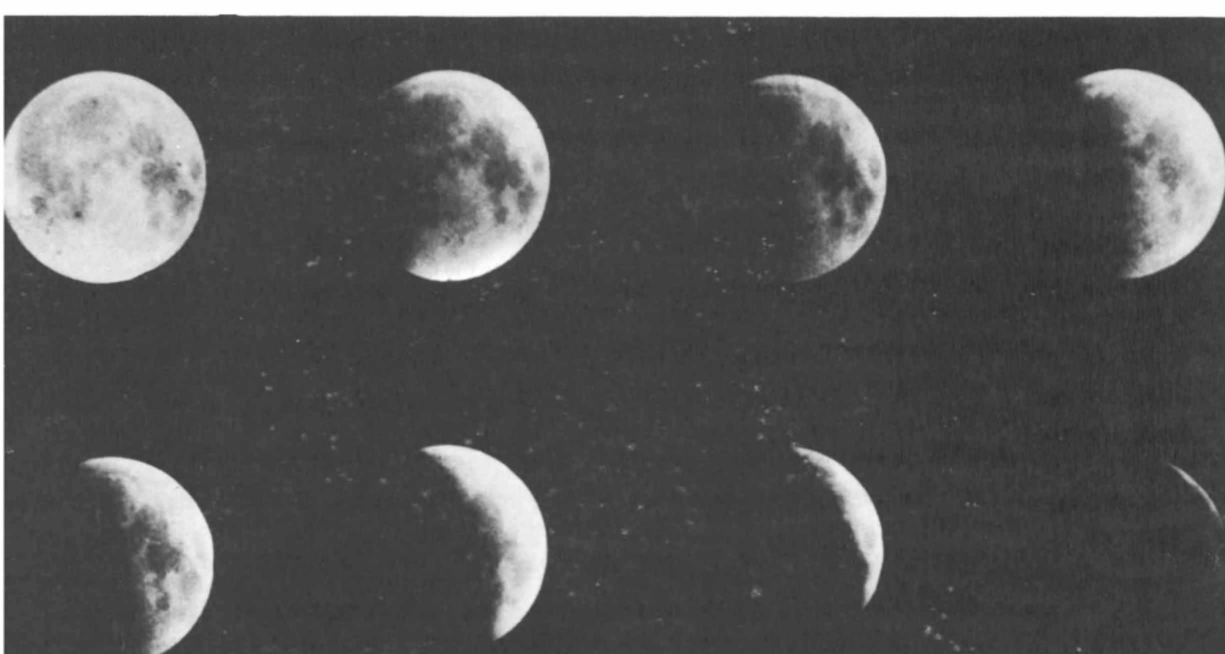
How Astronomical Discoveries Are and Aren't Done

No scientific fact or theory gains a definitive, general recognition before it has been tested by different scientists. Science differs from idle speculation in that its results can be tested by anyone if he or she so desires. True, there are two difficulties here. Not infrequently some phenomenon can only be observed using sophisticated, expensive apparatus, and nothing can be done without them. In other cases, in order to be able to check some calculation you will need adequate scientific training. You may need to know, say, other established facts from astronomy, mechanics, chemistry, or know the laws of physics, and have a proper mathematical background.

In this book we will try and introduce enquiring minds to how astronomers have been led to some conclusions, but this will require some effort on your part for at times you will be called upon to analyze a drawing or ponder upon some problem. But you will have to accept this because sciences based on mathematics and physics cannot, of course, be absorbed as easily as some science of descriptive nature. We hope the scholars who study such sciences will not be offended at this contrast. In any case they will agree that their descriptions and conclusions are clearer because they deal with developments more evident and familiar to a man in the street and do not require computation.

Many people believe that astronomical discoveries are made thus: an astronomer is sitting at his telescope and suddenly sees in it a new celestial body with some unusual movements and properties. True, occasionally

The lunar eclipse of 17/18 November 1956.



this happens, and often as a result of special search, they find a new nova explosion or a comet approaching the Earth from the depths of space. This is often done in the laboratory from pictures of the sky. But the method is not fruitful, and the major results come from the systematic, long-term study of the objects concerned.

By way of example, consider the discovery of a "high-velocity" star. Such a discovery may or may not be made through systematic examination of the apparent motion of stars in the sky. There are two ways here: one is to compile, as a result of many years of telescope observations, a catalogue of stellar coordinates with the maximum precision possible. On clear moonless nights mark (up to a hundredth of a second) the moment the star passes through the vertical line in the field of telescope and then climb up a ladder in the dark in order to read the angle between the telescope direction and the horizon on several microscopes. Several microscopes are required to reduce the error due to irregularities of the real circle divided into degrees and minutes, and due to tiny inaccuracy of dividing machine. The exact moment is fixed by a mark on a continuous tape on which a highly accurate clockwork located in the basement marks seconds as lines. The moment required is defined by measuring, through the microscope, of the exact position of the mark on the tape. But there is some uncertainty in the clockwork indication which floats, too. It is this error that is continually determined from stellar observations.

In the long run, a catalogue of coordinates of thousands of stars is compiled. Now the task is to compare it with a catalogue that was compiled *with the same accuracy* decades ago in order to identify stars with noticeably changed coordinates, the so-called high-velocity stars. The catalogues may have been compiled by people who are no more and they could not then make use of their catalogues for this purpose. Thus, astronomy to a large measure draws on the past and works for the future.

A simpler way is to compare pictures, rather than catalogues, of the same region of the sky obtained using *the same telescope*. But here, too, the time span between the pictures must be several decades.

Any discovery to some extent or other relies on the efforts of other people, on their successes and failures and their thoughts. Nowadays, to make a discovery one generally has to study much and know much. Observational astronomy requires sophisticated astronomical equipment, and theoretical astronomy a profound knowledge of physics and mathematics.

Science has been popularized in so many books which seek to tell the reader, in as simple and concise a way as possible, about complicated things that are sometimes too difficult to grasp. Perhaps, they give an impression that a scientific result is obtained just as simply as it is described, a bit of theorizing over a cup of tea being enough. The same impression may be produced by the stories that concerning some grey areas of science some scientist has one opinion, or another has another opinion, and yet another has yet another opinion. Although such differences are not uncommon, they are still supported by scientific evidence. You should not think that anyone can pronounce an opinion that would have some value for science.

The World of Solid Matter

We on Earth are most intimately acquainted with matter as a solid, but in outer space, it is far less often encountered as a solid than as gas. Solid matter, however, is structurally more complex and life requires "solid soil" for support.

Planets have a solid core (possibly not all of them and not always), the small heads of tremendous gaseous comets and small heavenly bodies (down to dust motes) are solids carried about in outer space.

Let us now look at them in more detail.

1. The Main Members of the Solar System

Distant Worlds – Satellites of the Sun

The Earth is a satellite of the Sun in space, whirling endlessly around the source of heat and light that makes life possible on Earth. Also circling the Sun are other satellites, the planets of the solar system. Each receives its share of solar heat and light depending on its distance from the Sun, and they are arranged in the following order: Mercury, Venus, the Earth, Mars, Jupiter, Saturn, Uranus, Neptune, and Pluto. We are only one fortieth as far from the Sun as Pluto and we are 2.5 times as far as Mercury. In almost every book on astronomy there is a graphic model of the solar system representing the Sun and the planets as various kinds of fruit, and their orbits – the paths of the planets around the central body – by different size circles. There is no need to describe this sort of model here again. Its chief purpose is to show the comparative sizes of the planets and the Sun, so illustrating graphically the enormous distances between the planets compared with their size.

We will merely mention that if the 149,600,000 kilometres that separate the Earth from the Sun – the *astronomical unit* of distance – are represented as one metre, the Sun will be the size of a cherry; the Earth a speck of dust less than a tenth of a millimetre in size; Jupiter, the largest planet, a pin head; and the smallest planets, Mercury and Pluto, specks of dust from one-half to one-third the size of the particle representing the Earth. They will not even be visible to the eye. In addition to the major members of the solar system, there are a great number of minor planets and comets, large and small, which will be discussed later. The satellites of the planets, including the Moon, which escorts the Earth and obligingly illuminates it at night, are also lesser members of the solar system.

In this chapter we will be looking at the major planets, but later in the book we will have much to say about the minor planets.

The course of the planets around the Sun and of the satellites around their planets is nearly uniform motion along a circle, the deviation being only a small one, Kepler found this three centuries ago when he refined the great discovery of the Polish scientific genius Nicolaus Copernicus (1473-1543).

Kepler's Laws

Kepler discovered three laws of planetary motion that completely define this motion. In his first law he showed that the planets go around the Sun in ellipses, one of the two foci of which always coincides with the Sun*. In his second law Kepler says that as a planet moves the line segment joining the planet and the Sun always covers equal areas in a unit of time. In his third law Kepler established that the squares of the periods P of revolution of the planets are proportional to the cubes of their mean distances from the Sun a, i. e.

$$\frac{P_1^2}{P_2^2} = \frac{a_1^3}{a_2^3}$$

where the subscripts 1 and 2 refer to any two planets.

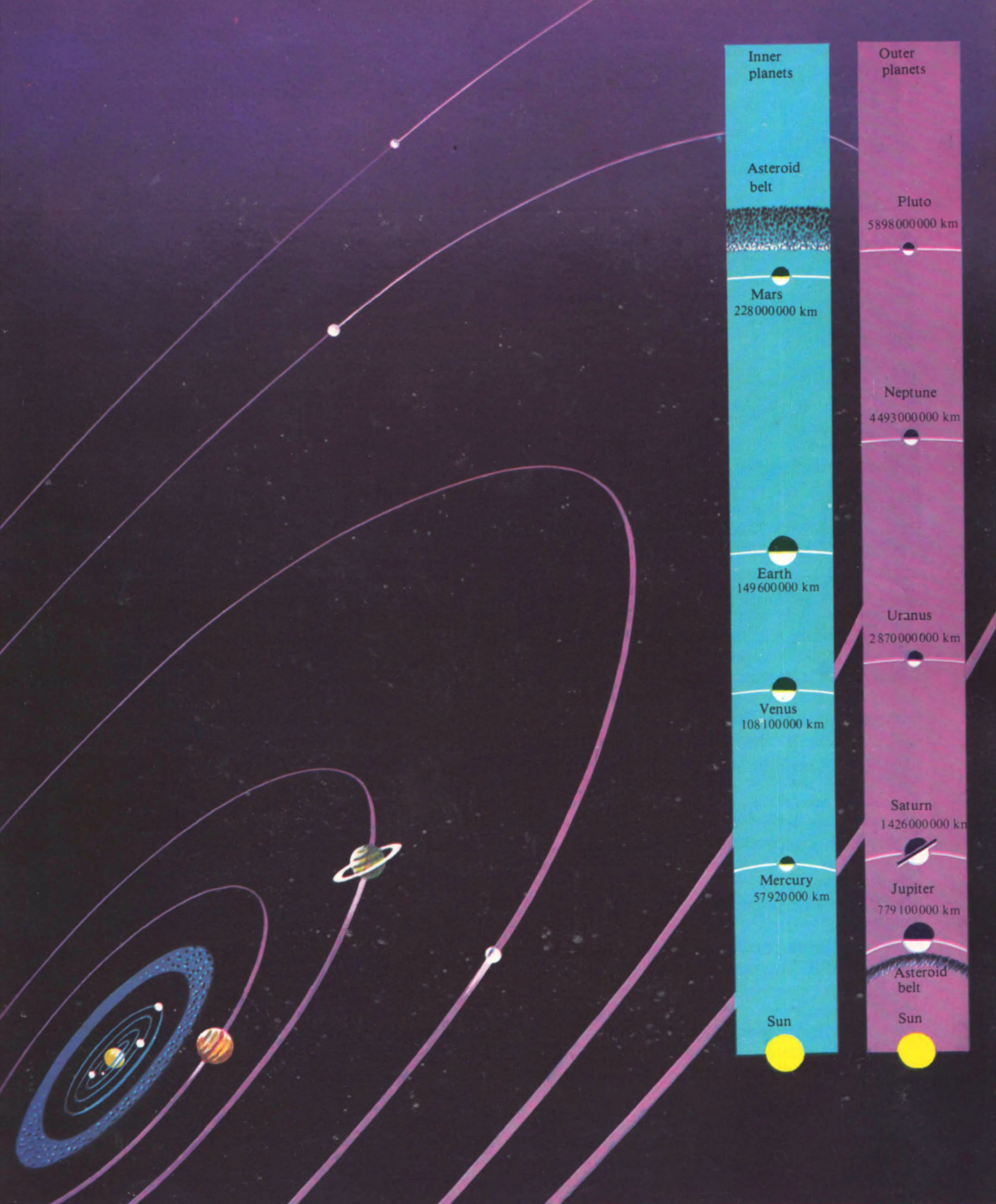
All the three laws were shown by Newton to be consequences of universal gravitation. They also hold true for the motion of satellites around their planets, and for the motion of any body attracted by gravity towards another. In some cases bodies may move not along ellipses, but along other, no longer closed, curves – the parabola or hyperbola. These curves have foci, too, and the main body will always be at the focus of the orbit. If the body is not moving in a closed orbit, we can no longer speak about a period of revolution and the third law becomes irrelevant, of course. Newton showed that for the elliptical motion a more correct version of the third law is

$$\frac{P_1^2(M_1 + m_1)}{P_2^2(M_2 + m_2)} = \frac{a_1^3}{a_2^3}.$$

It can be applied to any two masses M_1 and M_2, of which the first has a satellite with mass m_1, revolving around it with a period P_1 at a mean distance a_1, and the second mass M_2 has a satellite of mass m_2 with a period of revolution P_2 at a mean distance a_2. Using this formula we can compare, for example, the motion of the Moon around the Earth with the motion of the Earth around the Sun or with the motion of Neptune's satellite around its planet. If the masses of the satellites

* An ellipse is a curve in which the sum of the distances to it from two given points – called the foci of the ellipse – is constant.

The solar system. Shown on the right are the mean solar distances.



Inner planets
Asteroid belt
Mars
228 000 000 km
Earth
149 600 000 km
Venus
108 100 000 km
Mercury
57 920 000 km
Sun
Outer planets
Pluto
5 898 000 000 km
Neptune
4 493 000 000 km
Uranus
2 870 000 000 km
Saturn
1 426 000 000 km
Jupiter
779 100 000 km
Asteroid belt
Sun

are negligible as compared with those of the planets, they may be discarded in the above relationship. Then, when we apply it to two planets – satellites of the Sun – we may cancel the solar mass with the result that the formula takes the form in which Kepler himself gave it. His formula is approximate but precise enough for the planets of the solar system, since their combined mass only accounts for $^{1}/_{750}$ of the Sun's mass. The refinement due to Newton is exceedingly important in that it makes it possible to determine the masses of the bodies that enter Kepler's third law.

The gravitational pull on the planets due to each other is small as compared with the attraction by the Sun, but it causes deviations in their motion and slightly alters the shape and position of their orbits. These deviations are called *perturbations*. It is possible to predict these perturbations many years in advance if we know the masses of the bodies concerned and their orbits at a given time.

We can easily imagine the motion of a planet if we know the shape and position of its orbit in space, as well as the planet's position in the orbit at some instant of time. There are six quantities describing these data and they are called the *orbital elements*. For our purposes four of them would be enough.

Orbital Elements

The size of the orbit is given by the *semimajor axis* of the ellipse, a, expressed in astronomical units. The place where a planet is closest to the Sun and moves the fastest is called the *perihelion*, its antipode being called the *aphelion*.

The more elongated the ellipse, the greater the solar distances of perihelion and aphelion, and the farther away is the focus from the centre. The degree of elongation of the ellipse is described by the *eccentricity*, e, which is the ratio of the distance between the focus and the centre to the length of the semimajor axis.

For a circle $e = 0$, and as e tends to unity the ellipse's centre recedes to infinity. In other words, the ellipse stretches indefinitely, so that its branches tend to become parallel to each other to yield an open curve called a *parabola*. A still more open curve is called the *hyperbola*, it has e larger than 1.

The distance from the Sun to the perihelion is called the *perihelic distance*.

The third element, i, is the angle at which the orbital plane of a body is inclined to the plane of the Earth's orbit (ecliptic); it is called the *inclination*. For the planets, all of which orbit the Sun in the same direction, the orbital inclinations are very small. If the inclination is more than 90° (for instance, for some comets), it means that the body revolves in an opposite direction to that of the planets.

As the fourth element we will take some instant when the body passes through perihelion. We will denote it by T.

With the larger planets eccentricities are not large: Pluto has the greatest (0.249), followed by Mercury (0.206). The eccentricity of the Earth's orbit is only 0.017, or $^{1}/_{60}$, and if we draw the terrestrial orbit with semimajor axis 1 metre, the semiminor axis will be only $^{1}/_{7}$ millimetre shorter.

Space Near Our Home Planet

The boundary beyond which space begins is understood differently by different astronomers. Some of them believe that space begins at as low as 150-200 kilometres, others that it lies beyond the terrestrial atmosphere. But where this "beyond" begins – nobody knows, since the atmosphere thins away with altitude and gradually fades into the interplanetary medium. The everyday idea of interplanetary space is that it is a complete vacuum, emptiness. But apart from radiations (light, thermal and others) it contains gas particles, electrons, and cosmic dust. The density of these particles is now measured at various distances from the Earth and the Sun using satellite-borne devices. Several decades of space rocketry have substantially increased our understanding of the terrestrial neighbourhood. A detailed description of the Earth's atmosphere and methods of its investigations are beyond the scope of this book. We only need some idea for comparison with other planets, and so we will confine ourselves to some basic facts.

Ordinary clouds consisting of water droplets or small ice crystals are concentrated within the lower, convective stratum of the atmosphere, lying at altitudes between 8 and 15 kilometres depending on conditions. Noctilucent clouds, which by the way are a rather rewarding target of amateur observations, float at about 80 kilometres up. Aurorae polaris, which are essentially electric luminescence of air bombarded by

cosmic rays, at times reach altitudes of hundreds of kilometres. The density of the upper atmosphere varies markedly depending on active processes occurring on the solar surface, and also varies diurnally. At 1,500 kilometres it is on average about $5 \cdot 10^{-18}$ gramme cubic centimentre. Some traces of the atmosphere are to be found up to 3,000 kilometres.

Planetary probes and satellites have also measured the density of interplanetary dust at large distances from the Earth. Its properties can also be estimated from how it scatters sunlight, making the day sky brighter. So far little is known about particle sizes and the concentration of the dust. It may well be that the Earth with its atmosphere hampers the motion of dust particles and catches some of them. Some of these particles may have formed in explosions that occur when meteorites hit the lunar surface.

Thus, a wide variety of species of gas particles, molecules and atoms hurtle through interplanetary space. In addition, we find there large meteorites, and smaller meteoric bodies down to specks of cosmic dust. But this is not the whole of the story for wandering through the space there are electrically charged particles, protons and electrons, which also behave interestingly. Much of the difference between them stems from their different kinetic energy.

But the most striking development was the discovery of the Earth's radiation belts. Satellite-borne instruments have found a flattened "cloud" of electrically charged particles. It lies near the plane of the magnetic equator of the Earth. There are ring zones inside the "cloud" that are regions with higher concentration of particles. This electric terrestrial crown stretches up eight Earth's radii (up to 50,000 kilometres). The highest concentration in the inner belt is at an altitude of about 10,000 kilometres, the major particle species being protons. The outer belt is wider and contains electrons with energies up to 100,000 electron-volts.

The inner belt forms as a result of disintegration of atoms in the atmosphere bombarded by cosmic rays. The latter consist of particles moving at enormous speed, which makes them powerful ionizing agents. They come from the Sun and other more distant regions of space, their origins are not yet completely clear. The fragments of the disintegrated atoms are electrically charged and begin to move in the terrestrial magnetic field spiralling along its lines of force.

UV picture of Venus from 720,000 kilometres (*Mariner 10* on 6 February 1974) showing cloud layers. The latter are seen to have a distinct orientation despite the fact that the planet is rotating rather slowly.

So the Earth's magnetic field is a trap for such particles where they accumulate.

The origin of the outer radiation belt discovered by the earlier Soviet satellites is also not clear as yet. Some space accelerator imparts great energies to particles that at times come to us during solar flares.

Although of huge scientific interest, the radiation belts are a hazard for astronauts. So the investigation of the behaviour of the radiation belts is of significance for the development of aeronautics.

Planetary and Lunar Studies

The planets are distant earths, sisters (or cousins if you wish) of our Earth. Still these distant earths are our nearest neighbours in the endless wastes of the Universe. Through the telescope we see even their disks and, say, Jupiter at a 50-fold magnification is just about the same size as the Moon to the naked eye. Nevertheless, each planet harbours many puzzles that still await answers, and alas we know less about them than we do about many immeasurably more distant stars. The planets only reflect sunlight, therefore their spectrum essentially coincides with that of the Sun. This finding is the major recent addition to our understanding of the nature of planets. More information

comes from radio observations of the planets and the Moon and also from planetary probes.

It was not until recently that astronomers understood that they do not see the surfaces of many planets, but only ever-changing clouds that cover and blanket them from us. Such is the case with Venus, Jupiter, Saturn, Uranus, and Neptune. Although the presence of clouds suggests the presence of thick atmospheres on these planets, especially on the last four, it adds virtually nothing to our knowledge of what their surfaces look like. These clouds, like the yashmak of a Muslim woman, shield from us the face of many planets.

We can now measure the temperature of a planet by means of thermoelements and other devices, and also from their radio radiation. These temperatures refer to the visible sides of the planets, i. e. in some cases to the surface proper, and in the other cases only to a particular strata of their respective atmospheres. Data on the temperatures of the planets sometimes make us reconsider to a significant measure our earlier notions of their physical nature. Its atmosphere to a large extent predetermines the temperature conditions on a planet. The life of a planet is harder if it has no atmospheric "blanket".

On Earth the clouds and the air itself protect the soil from getting too hot during the day, and at night they check the release of the accumulated heat into space. Day and night temperatures are thus fairly balanced. It is clear, too, that the levelling out of the surface temperature is helped by the planet's rotation about its axis with respect to the Sun, and that the more rapid this rotation is, the greater its effect.

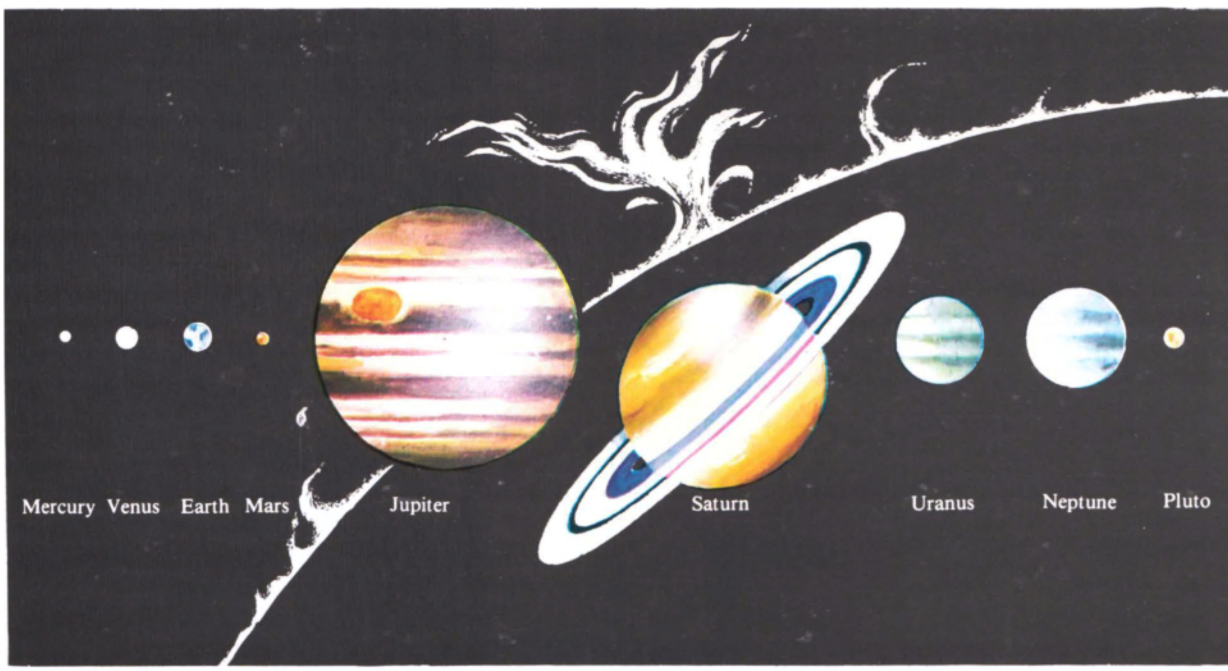
Spectroscopic analysis cannot give us as much information about the planets as it can about the stars because the planets shine by the reflected light of the Sun. It would be wrong, however, to think that it cannot help at all in studying the planets. It was long ago suspected that determining, using the Doppler effect, the velocities of two opposite edges of a planet with respect to us would give the period of rotation of the planet about its axis. In this way the periods were determined for Uranus and Neptune.

The energy distribution in the continuous spectrum of a planet is not a carbon copy of the solar spectrum. While a planet, as is often and imprecisely stated, merely reflects sunlight like a mirror, this mirror distorts. Indeed, the energy distribution in a planet's spectrum is not the same as it is in the solar spectrum, because the surface of any planet, like that of any body, is not an ideal mirror and is not an ideally white surface, it therefore does not reflect different wavelengths uniformly. In general, red surfaces reflect red rays the best and so the red portion of the spectrum reflected by them will be brighter than the rest of the spectrum compared to the spectrum of a source illuminating the surfaces. It is this greater brightness of red rays in their spectrum that is responsible for the red colour of their surfaces.

It has long been known that gases consist of molecules moving randomly with arbitrary velocities, their average velocity depending on their masses and temperature. The smaller the mass of the molecules and the higher the temperature, the greater the average velocity. On the other hand, when a gas particle reaches a certain limiting, or *critical*, velocity, a planet is no longer able to hold it and stop it from flying off into the airless interplanetary space. Knowing the force of gravity at the surface of a planet (increasing with a planet's mass and decreasing with its diameter), it is possible to calculate this critical velocity. For the Earth it is 11.2 kilometres per second and for the Moon 2.4 kilometres per second. The rate at which the atmosphere of each planet is dissipated has been calculated, and it was found that if the Moon and Mercury had at one time had atmospheres they must have been dissipated very rapidly. This explains why we do not observe an atmosphere on these celestial bodies at the present time. The molecules of their respective atmospheres have long since left their weak hosts, the Moon and Mercury *. The presence of an atmosphere almost as dense as the Earth's on our beautiful neighbour Venus was first established from observations made in 1761 by the Russian scientific genius Mikhail Lomonosov (1711-1765). It has been theoretically predicted and confirmed by probes that the Martian atmosphere is thinner than ours.

The larger planets have very extensive atmospheres. The gravitational attraction of these planets is capable of holding (the more so since their surface temperatures are low) even the lightest gases, such as hy-

* Using instruments in orbit it has been found that Mercury and the Moon possess traces of atmospheres, which, however, are so thin that are not in any way comparable with the tangible atmosphere of our Earth.



The relative sizes of the nine planets and the Sun shown to scale.

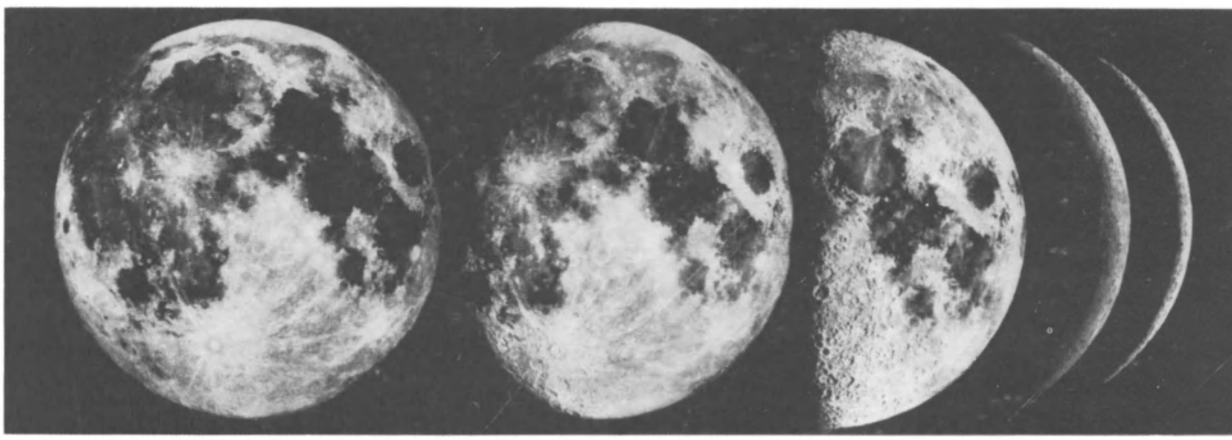
drogen, which have the highest average molecular velocities. Light gases easily escape from the terrestrial atmosphere. Individual molecules escaping from the atmospheres of Jupiter and Saturn are so few in number that their loss has not yet shown up to any marked degree.

The spectra of planets with atmospheres differ from the solar spectrum in more than just the energy distribution. The atmosphere of a planet superimposes on the Sun's spectrum its own make-up, as it were. It produces new dark lines and bands. The same occurs in the Earth's atmosphere. In fact in passing through the latter the sunlight is absorbed by the constituent gases, which results in the dark lines in the solar spectrum peculiar to these gases. The solar spectrum, as observed from the Earth's surface, contains lines belonging to atmospheric water vapour, oxygen, and nitrogen. These lines, termed *telluric*, can be distinguished from those belonging to the Sun because they become stronger as the Sun approaches the horizon and the thickness of the atmosphere through which its rays must pass increases. Another technique is based on the fact that the lines in the spectrum from the Sun's edge are Doppler-shifted due to rotation, while the telluric lines remain in their normal positions.

Sunlight penetrates the atmosphere of a planet and on being reflected from its surface passes again through the planet's atmosphere before reaching the Earth. The additional absorption of sunlight by the molecules in the planet's atmosphere causes either an intensification of the telluric lines as compared with the spectrum obtained directly from the Sun, or the appearance of new lines if that gas is absent in the terrestrial atmosphere.

The Old and New About Our Eternal Satellite

New evidence is habitually evaluated based on the old evidence. Thus, in order to outline the new discoveries, I should recapitulate the numerous and interesting facts we know about the Moon, our eternal satellite. But I hope, dear reader, that you know the basics. So you are sure to know that the Moon is four times smaller than the Earth in diameter and 81 times in



The phases of the Moon.

mass, and that its gravitational pull is six times weaker. If you were to jump on the Moon you would rise six times higher and would fall much slower than on the Earth. It is common knowledge now that the Moon always faces us with only one side, and until fairly recently nobody even hoped ever to find out what the farside looks like. True, there was plenty of speculation in science-fiction novels, and scientists have long recognized that it should be generally the same as the nearside. It should also be devoid of water, and atmosphere, and the mountains should also be in the forms of cirques and craters.

The diameters of lunar ring mountains are much larger, relative to the height of the encircling walls, than they would be if they were craters of terrestrial volcanoes.

The cherished dream of astronomers to see the farside of the Moon came true only two years after the first satellite was launched in the USSR. On 4 October 1959 the lunar probe *Luna* 3 looped the Moon and took picture of the farside for 40 minutes. It was no less remarkable that the images of these automatically developed pictures were sent back to Earth by television.

Even now the performance seems striking, although ever since radio and TV information (even of higher quality) has been sent back from as far as Jupiter, i. e. more than 600,000,000 kilometres. Such are the fabulous strides made by radio technology.

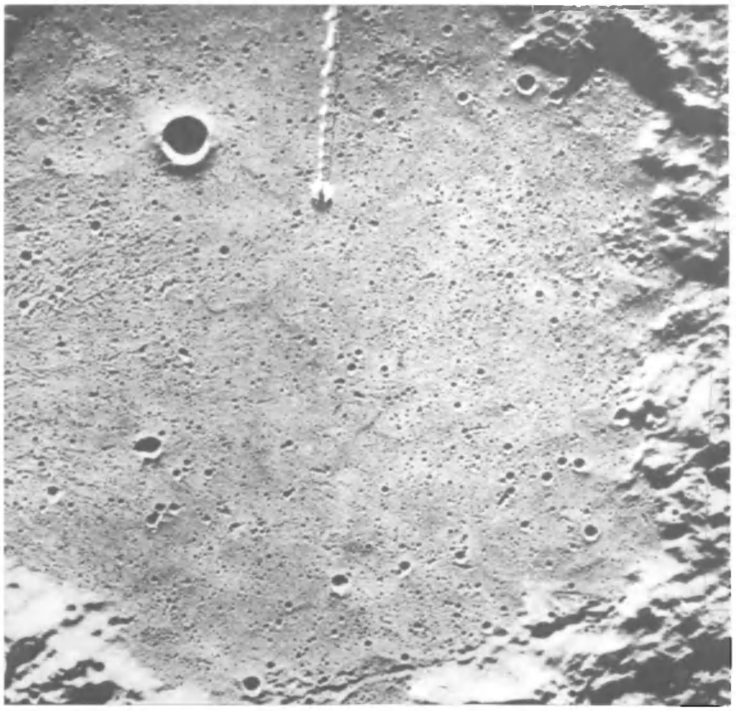
To obtain pictures of the farside the cameras in orbit had to be trained accurately at the right time at the Moon. It was necessary to orient the probe in space, since for a number of reasons it always tends to rotate. This had been achieved by a suitable system.

It was also necessary to take pictures of some of the nearside so that this information could be used to bridge the new features to the old ones when it came to charting the farside.

Later on, the Moon was attacked by apparatus that banged into its visible side having sent back some details of the approaching lunar panorama. Then lunar probes began to circle the Moon and take pictures from different angles and distances. This enabled maps of both sides to be made with more detail and precision than some maps of some terrestrial regions. It turned out that the farside has less "maria" than the nearside (maria are the low-lying regions with dark surface flooded by erupted lava and ejecta, *basalts*), and it is as heavily cratered as the nearside. A new feature found on the farside is the large circular depressions with light bottoms called *talassoids* (mare-like). Lunar craters have been named by agreement between international scientists. On the nearside there are the craters Plato, Archimedes, Copernicus, and so on whilst on the farside we encounter more contemporary names like Moscow Sea, craters Korolev, Tsiolkovsky, Jules Verne, Edison, Herzsprung, Mendeleyev, Gagarin, etc., etc.

Just like the terrestrial surface, the lunar crust was

The view of the Earth from a lunar orbit (*Zond* 7, 1969).



The Ptolemaeus crater, 150 kilometres in diameter. The bottom of the crater (light) is relatively flat. Seen at the top is the arm with a gamma-spectrometer at *Apollo 16* designed to measure the radioactivity of the lunar rock.

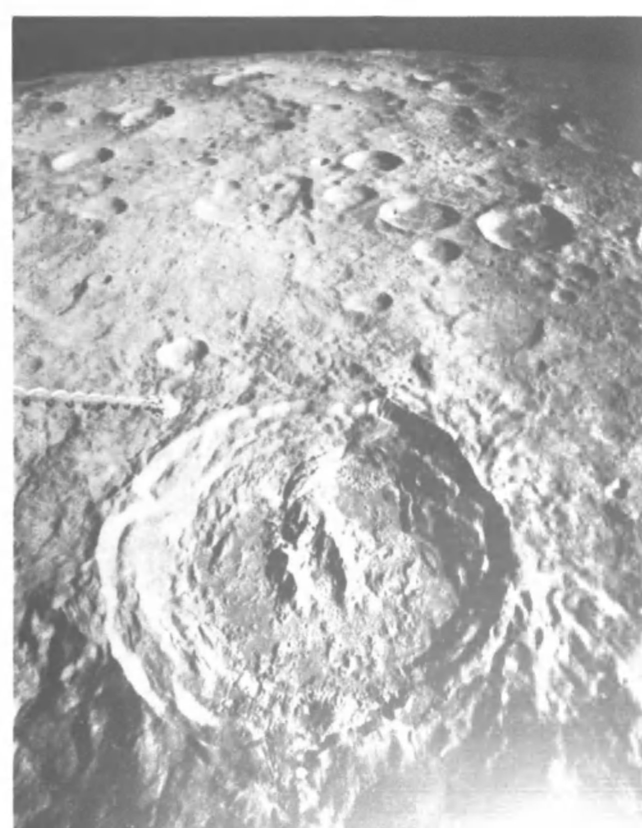
formed by alternating orogenesis and subsidence of vast regions flooded by melted stony mass (magma) from within the Moon, which is like the lava ejected in volcanic eruptions. The lunar crust has wrinkle ridges. Many of the major lunar craters are assumed to be of volcanic origin. So from the Tycho crater, 60 kilometres in diameter (named after the Danish astronomer Tycho Brahe) in the southern lunar hemisphere radial rays extend nearly up to the middle of the nearside. It was a thin layer of material ejected in the explosion that caused the crater. In the neighbourhood clusters of smaller craters are found. Lunar craters resemble the lava lakes in Hawaii, which are smaller and have lower rims. It appears that the lunar craters do not resemble conventional volcanoes in the least, although some of them have steep cones at the centre more like volcanoes.

The Moon is considered to be a dead world without any signs of life or even geological evolution (it would be more correct to call it selenological evolution from the Greek for the Moon "Selena"). Of especial interest therefore are the findings of N. A. Kozyrev of Pulkovo who succeeded, when he was photographing the spectrum of the lunar cirque Alphonsus, in observing a phenomenon that was similar to a volcanic eruption. He found in 1958 that the central hill of this crater got red and its spectrum showed carbon lines. It seems that some carbon vapour erupted into free space around the Moon, and the vapour glowed in sunlight just as a comet's head does. For the Moon such developments are undeniably extremely rare and hard to hit upon.

The smaller the craters, the more there are. So there are about 200 craters 20 to 50 kilometres across and more than 600 from 5 to 20 kilometres across per ten craters of diameter greater than 100 kilometres. The finest details distinguished through the largest telescopes are slightly less than 1 kilometre in diameter. In some places average-size craters are found to lie in the form of chains up to 600 kilometres in length. On Earth, too, along the so-called fracture zones in the crust, volcanoes are arranged in chains, but they are few and far between.

When sunlight falls obliquely on the lunar surface the mountains cast long shadows. Striking view opens up before a telescope observer who sees the new sun as if setting the mountain peaks on fire and 8-km mountains gradually "growing out" of the darkness. The

View of a 75-kilometre crater. Note the unusual form of the central part.



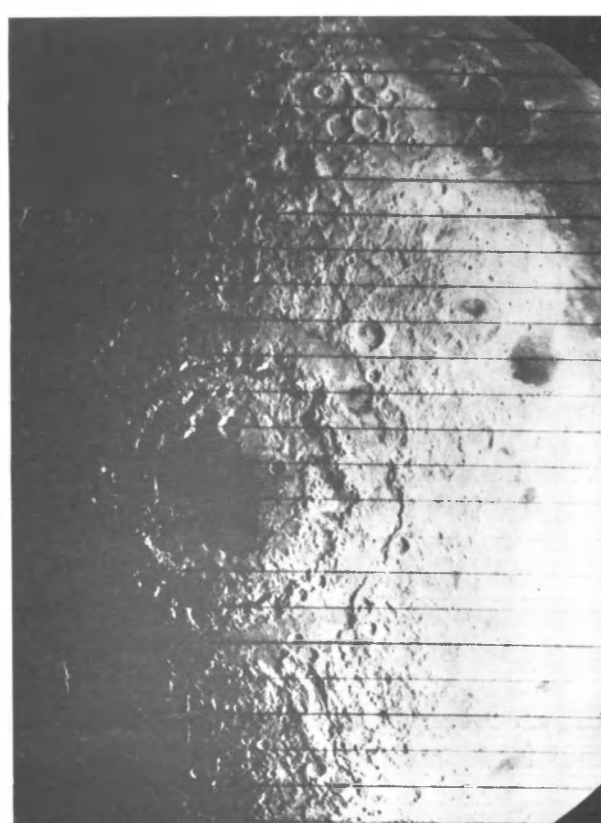
Part of the Mare Orientale (*Orbiter 4*), one of the most noteworthy features of the Moon's surface. Note the "crevasses" and crater chains several hundreds of kilometres from the sea's centre.

Moon is not only exceedingly mountainous but it is strewn with sharp rubble and boulders, so that it would give a rough ride to any vehicle. But this is to be expected since the Moon has neither air nor water to weather away its mountains and rocks. On the other hand, plenty of cosmic dust particles must have precipitated on the Moon, just as they have on the Earth, but when they enter our atmosphere, the dust particles evaporate. Lunar craters come in an infinite variety of sizes and the smallest of them, about 0.5 metre in diameter, cover the whole of the surface even at the bottom of the crater Alphonsus, even though through the telescope this crater's bottom seems to be smooth as compared with the mountainous regions.

It was found that small craters, about 10 metres in diameter, often have no high rims steep on the inside, or the rim is only several metres high. Both on the inside and on the outside the slopes are gentle, as if eroded by water. But there is no water on the Moon and so the only explanation is that ancient craters have been ironed down by meteoritic impacts, which are not cushioned by the action of an atmosphere as is the case on Earth. For millions of years, small meteorites have blunted the edges of ancient craters. Later and smaller craters are less eroded, and their walls are steeper, whereas some of the older ones are now just flat low-lying sites. To counterbalance the volcanic hypothesis of the formation of lunar craters a new hypothesis has long been put forward that the ancient lunar craters are due to the impacts of large meteorites, perhaps like those that according to some hypotheses collided and merged to form the planets themselves. Colliding with the surface of the Moon at a relative velocity of several kilometres per second, the meteorites must explode like bombs producing craters that are much larger than the meteorites themselves.

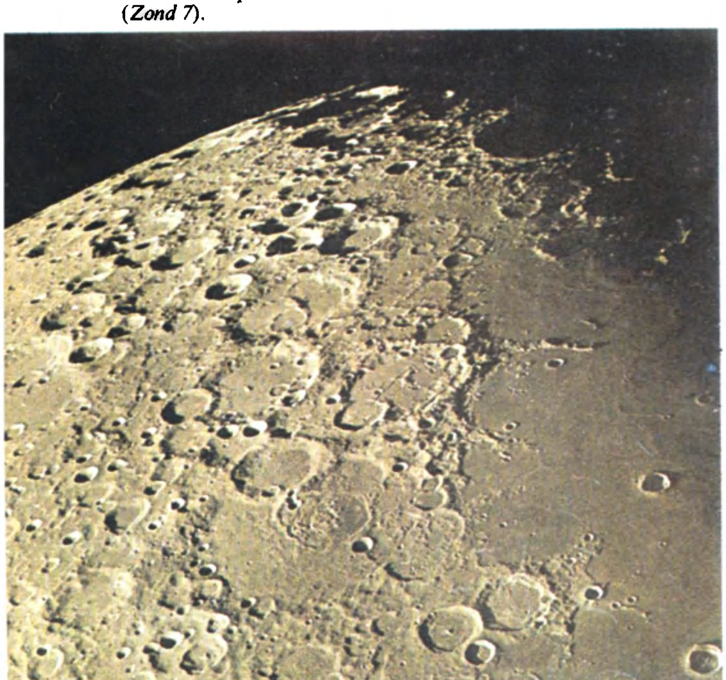
The evidence of the destruction of ancient craters by meteorites does not account for the patterns formed by lunar features. They cannot be explained by random falls of meteorites, but rather they suggest some connection with activity underneath the Moon's crust. We will look at this issue in more detail later in the book. The many lunar "cracks", which were discovered long ago by telescope, are actually long valleys less than 1 kilometre wide and with very *gentle* slopes.

Mare Orientale at the edge of the Moon's farside. Note the seven concentric mountain chains, cracks and crater chains which are up to 700 kilometres away from the centre.



The Moon's surface from 10,000 kilometres (*Zond 7*).

A moonscape–the Ocean of Storms and numerous craters (*Zond 7*).

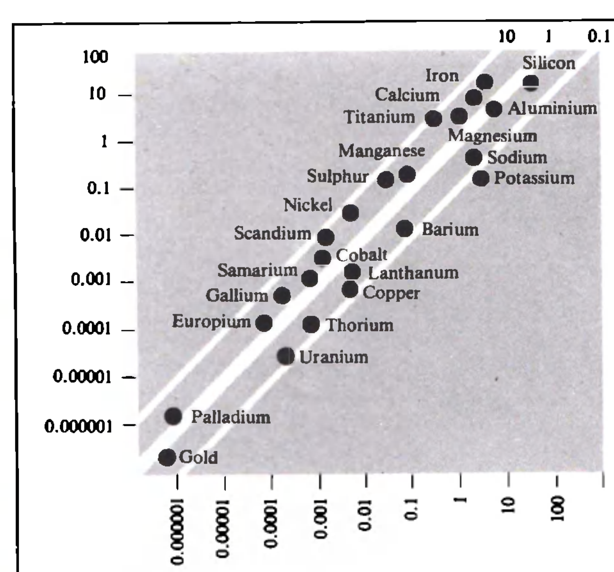


These are the edges of old cracks ironed out by meteoritic erosion and filled with rubble. On the other hand, in places the "cracks" are a chain of merged small craters. These craters could be of volcanic origin, since it was easier for lava to find its way up along the cracks.

My personal opinion is that the circular maria and, at least the larger craters, are caused by volcanic eruptions through the lunar crust, whereas meteoritic impacts have only been smoothing the relief producing for the most part small craters. In our time, too, in the vicinity of the Earth meteorites are the rarer, the larger they are.

And what can be said about the lunar dust? In 1966 the Soviet probe *Luna 9* was the first to softland on the Moon with its hardware in perfect order. It transmitted to the Earth the first magnificent panoramas of moonscape, and enabled the strength of lunar soil to be estimated. The soil appeared to be tough enough for the lander not to sink noticeably though. Between the second and third transmission sessions the probe shifted a bit. Thus, the ground under it yielded but no thick layer of dust was found.

The abundance of chemical elements in Moon rock (ordinate) relative to the Earth's crust (abscissa), the wide, white line corresponding to equal abundances.



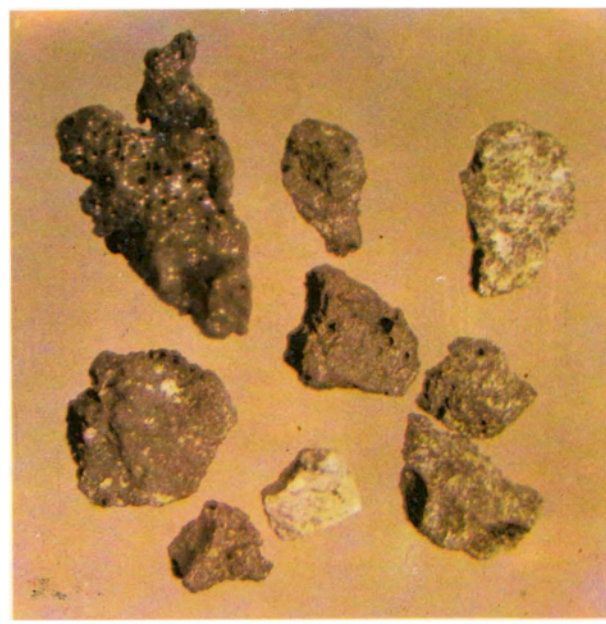
The surface layer of the Moon, called the *regolith*, has a porous structure. It was formed from a meteorite-crushed initial layer mixed with the meteorites and compressed in vacuum.

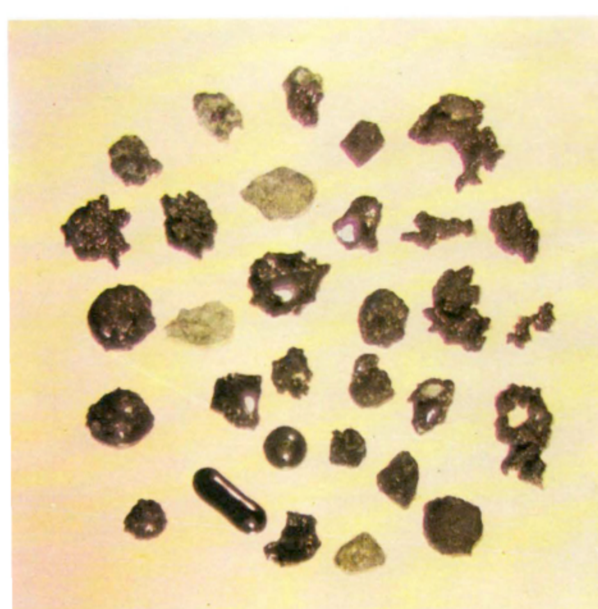
It would be difficult for us to adapt to very sharp variations of temperature on the Moon. The lunar day is two Earth-weeks long and during this time its surface at the equator heats up to 125°C. But during the two-week night the temperature drops down to − 150°C. On Earth our blanket of the atmosphere, like a tender nurse, protects the planet from excessive loss of heat to space, and from the scorching rays of the Sun. Astronomers, often curse the atmosphere as a nuisance, but we could not survive without it. The excessively small thermal conductivity of lunar rocks lies behind the fact that during a lunar eclipse, when the Moon is blotted out by the Earth's shadow, the Moon's surface temperature drops by 225°C in just an hour. This is because the surface layer of the Moon is a porous, pumicelike matter.

During the first few minutes heat is only lost from the outermost layer, several millimetres thick, because of the low thermal conductivity of lunar rocks. And it is only where the thermal conductivity is larger, that the supply of heat from the bulk prevents the material from cooling fast, with the result that the surface is for a while warmer there than in the surrounding areas.

Evidence derived from lunar eclipses using infrared detectors revealed several hundred "hot spots" nonuniformly scattered over the Moon. The temperature in them is 48°C higher than in the neighbourhood (in particular, higher than at the crater Tycho). Among them we find the craters Copernicus and Kepler, but they are mostly small craters several kilometres in diameter, which appear bright during full moons or have bright rays. They are even not as numerous on the terrae, as they are in the maria where craters are few. Many of these craters are young. Craters similar in appearance often differ markedly in heat release, the reason behind this difference has not yet been established.

Radio observations of V. A. Troitsky and coworkers of Gorky, the USSR, enabled us to look under the apparent surface of the Moon. Radio waves emitted by the Moon appeared to be less affected by temperature variations than infrared radiation, thus suggesting that radio waves come from deeper strata. This team of radioastronomers was led to conclude





Moon rock samples

that the temperature of lunar rock grows with depth and that the density of the material more than 4 centimetres down is about 1 gramme per cubic centimetre. This was confirmed later.

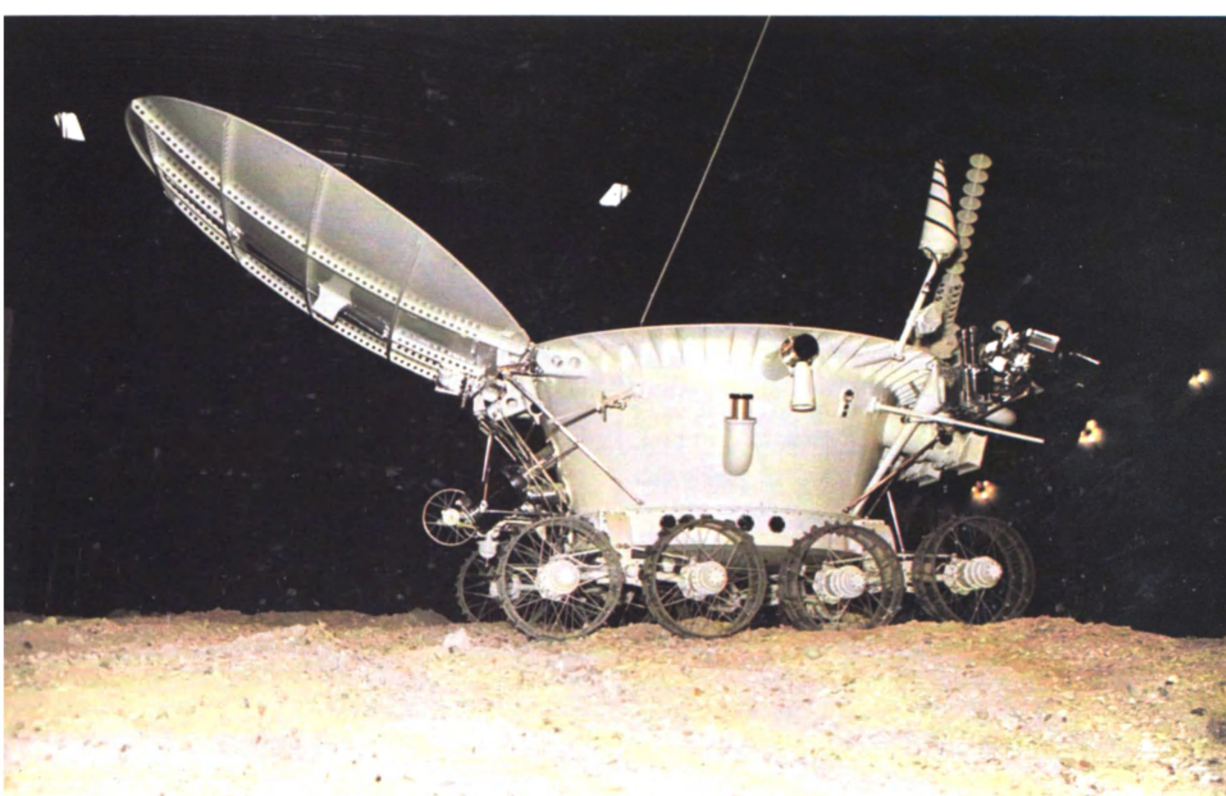
A Czech astronomer, Link, suggested that some areas on the Moon may luminesce under the action of the solar corpuscular radiation. Kopal and Rackham at the French high-altitude observatory Pic du Midi actually observed in 1963 fluorescence near the crater Kepler. During a two hours' period the brightness of this region in red light doubled twice. There was a solar flare some time before, and electrically charged particles ejected in this flare may have reached the Moon 8 hours later and bombarded the lunar surface to produce the extra luminescence there.

Astronomers were thus successful in penetrating even from Earth under the visible surface of the Moon. The Moon probes that have softlanded on the lunar surface have shown that it was possible for astronauts to land there, too.

Lunokhods and First Astronauts on the Moon

Who has not once been fascinated by the scientific thriller *The First Men on the Moon* by Herbert Wells, and also by other books about flights to the Moon. At long last the dream came true in a fantastically short time after the first Soviet sputnik defied the terrestrial pull. In July 1969, two American astronauts in *Apollo 11* landed on the Moon. For the first time the foot of man stepped on another celestial body. The event was

Lunokhod 2.



crucial in two respects. First, man had literally gone into the heavens. Second, the landing made it possible to install on the Moon various instruments, in particular seismographs, and perform direct experiments on our satellite. Ever since astronauts have softlanded on the Moon five times, walked on its surface, made observations, and returned safely back to Earth.

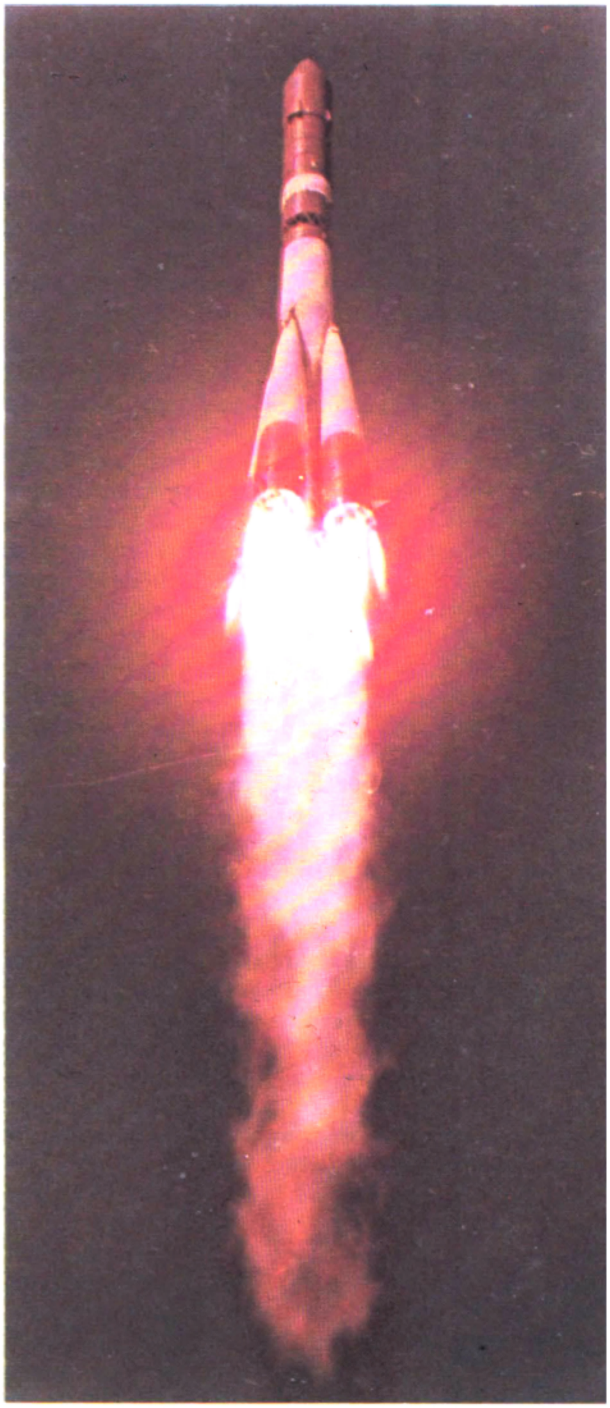
Altogether, the astronauts have spent about 300 hours on the Moon, of which time they have stayed 80 hours outside the ship in the airless space. This time has been spent in strolling over the lunar surface, taking pictures, and making movies, and collecting samples of lunar rocks.

Their strolls and moonscapes were televised back to Earth. The astronauts had space suits on to protect them from the vacuum and the lethal action of shortwave and corpuscular radiations of the Sun and other celestial bodies. They found one of the automatic devices that had softlanded earlier on the lunar surface and took its parts and rock samples back to Earth for thorough investigation. During the last landings the astronauts drove about the lunar surface in a moon car.

The lunar module landed with two astronauts, whereas the third one orbited around the Moon in the main module. Their mission on the lunar surface completed, the two astronauts started in a part of the lunar module, orbited for a while as a lunar satellite and then docked with the main block. The lunar module was then jettisoned and the astronauts returned on Earth in the main block.

Evidence from space flights has been accumulating at such a pace that any book is in danger of becoming dated as it goes into print. The interested reader is therefore referred to the periodicals for more details.

The Soviet Union has also made some advances with robot spacecraft. In September 1970 *Luna 16* landed on the Moon in the Sea of Fertility, drilled into the soil, and returned samples to Earth. In September 1972 the sister craft *Luna 20* obtained a core sample from a mountainous region where astronauts couldn't have got. Another Soviet robot, the Lunar Roving Vehicle *Lunokhod 1*, worked for 10.5 months (from November 17, 1970) in the Sea of Showers. Roving about on its unusual wheels, it examined 80,000 square



The start of a *Soyuz* launcher.

metres of lunar surface, studied the physico-mechanical properties of the lunar soil at hundreds of points, and performed chemical analyses of the soil at 25 points. The vehicle's TV cameras transmitted more than 200 panoramas and 20,000 pictures of the lunar surface. The French-built laser reflector installed on the top of *Lunokhod 1* allowed the distance to it to be precisely measured from the propagation time of a laser beam.

Lunokhod 2 softlanded in 1973 in the Sea of Serenity at the bottom of a 55-kilometre crater. It stayed there for 5 lunar days (4.5 months) and covered 37 kilometres, almost four times as much as Lunokhod 1. It also studied the surface features and properties of the lunar soil.

The evidence from these missions is vast and not all of it has yet been processed, discussed, and understood.

The soil samples enable the scientists to gain some insight into the chemical, mineralogical, and petrographical composition of the material at different sites on the Moon. On the one hand, we have direct evidence that there is no radical difference between the Earth and its satellite. We find the same chemical elements and minerals on both and in similar proportions, but there are also some differences between the two bodies, and differences from site to site on the Moon. The age of the lunar samples is from 3,100 to 4,200 million years, so in general the Moon and the Earth are about the same age (no less than 4,600 million years).

Judging from the samples, lunar soil is a loose material with a density of about 1 gramme per cubic centimetre, consisting of fragments of crystalline rock and granules of natural glass with some fine cement and fixed microscopical gas bubbles. This is a product of the erosion of the exposed surface of bedrock in a vacuum under the external impacts, such as meteorite bombardment, irradiation, and sharp temperature changes, which all have occurred during the unbelievably long periods of time. The regolith layer at various landing sites was from several metres to several tens of metres thick. The regolith has no distinct lower boundary and passes gradually into the underlying rocks. The surface layer of regolith contains the most heavily processed material, the middle layers contain many small fragments, and the lower layers contain large-size pieces of rubble. The lunar bedrocks, as the studies of the lunar samples showed, fall broadly into two types of igneous rock: the lunar maria are composed of basalt with a density of 3.2 grammes per cubic centimetre; and the terrae, of anorthosites.

The diurnal temperature variation on the regolith's surface is nearly 300°C but it becomes less marked with depth. So at a depth of several metres, the temperature is virtually constant and equal to the mean daily temperature. At larger depths the temperature is higher due to the influx of heat from the bulk of the planet that is due to decay of radioactive elements in the crust of the Moon. The Moon's crust is several tens of kilometres thick. Deeper still lies a material of another nature, and the centre of the Moon is a small tough (melted) core.

Direct information about radioactive elements on the Moon come from gamma-measurements first performed by *Luna 10*. The gamma-radiation was estimated to be nearly twice that of granites, the terrestrial rocks with the highest content of radioactive elements. But 90 per cent of the total lunar radiation is due to nuclear reactions occurring under the action of cosmic rays. The material of the mountainous regions of the Moon seems to be closer to the initial composition of its crust and to that of meteorites not so rich in radioactive elements. In certain places the initial crust sank later, the material melted, and the surface was flooded with basalt rocks.

It should be noted that gravitational measurements over different areas of the lunar surface showed that the pull of gravity above circular maria was about one per cent stronger than above adjoining regions. These gravitational anomalies must be caused by extra masses over these surfaces. These masses were called *mascons* (an abbreviation from *mass concentration*) though their origins are still discussed.

The layer of dust that has come from space or been produced by the disintegration of rocks is hardly noticeable in some places, whereas elsewhere it may lie in a thick layer. It was a nuisance when it deposited on the astronauts and their instruments.

The astronauts installed seismographs in five places on the Moon's surface to record lunar quakes, meteoritic impacts, and impacts from artificially produced explosions. This way the velocity and paths of the propagation of seismic waves are studied to obtain data on rock behaviour on the Moon. So they have once recorded the fall of a large meteorite.

It came a bit as a surprise that the lunar rocks are laminar in nature, the strata being several metres thick. This is an indication of the complicated processes that have been forming the rocks during long periods of time. It is now unquestionable that at least some of these features are due to impacts. Meteorite falls crushed the rock on the Moon, and the fragments spread and compacted.

We will now discuss what astronauts will see on the Moon in future if they can stay there longer. In their electrically heated space suits fitted with portable life support systems astronauts would leave the landing module through an air lock and step on the lunar soil in airless space. The Moon is four times smaller in diameter than the Earth, therefore its surface is more curved – the horizon is closer to the observer and he cannot see as far as on Earth. Even from the centre of a large crater the astronaut would only see the summits of the mountain ridge that forms the border of the crater.

Astronauts can only communicate on the Moon using walky-talkies, since sound waves do not travel in a vacuum. So if a huge meteorite hitted the surface producing a new crater, the astronauts would not hear it.

There are suggestions that the first base of a lunar mission be put underground to ensure adequate thermal and radiation protection. A roof will also be required on the Moon not to protect the astronauts from the rain (there is no rain on the Moon) but to protect them from micrometeorites. They can also come in for a sort of stone rain.

The temperature in the underground shelter (without heating) would be more stable, remaining close to 0°C. The heating could be done using a nuclear reactor and the cave, if pressurized, could be filled with air. A small amount of air might either be brought from the Earth or it could be obtained chemically from lunar rocks. It may well be that water as well could be extracted from the rock, although I think that this will be unlikely in the first missions.

When observing a sunrise astronauts will not see a dawn, and the first thing to appear in the dark sky will be the pearly rays of the solar corona, a halo around the Sun. On Earth it is only seen during total solar eclipses. Shadows on the Moon are black and sharp.

When a place gets in a shadow, the temperature drops drastically. We on Earth feel in a shadow the warmth of the heated air, although here too in the sunshine we get additional heat through direct radiation.

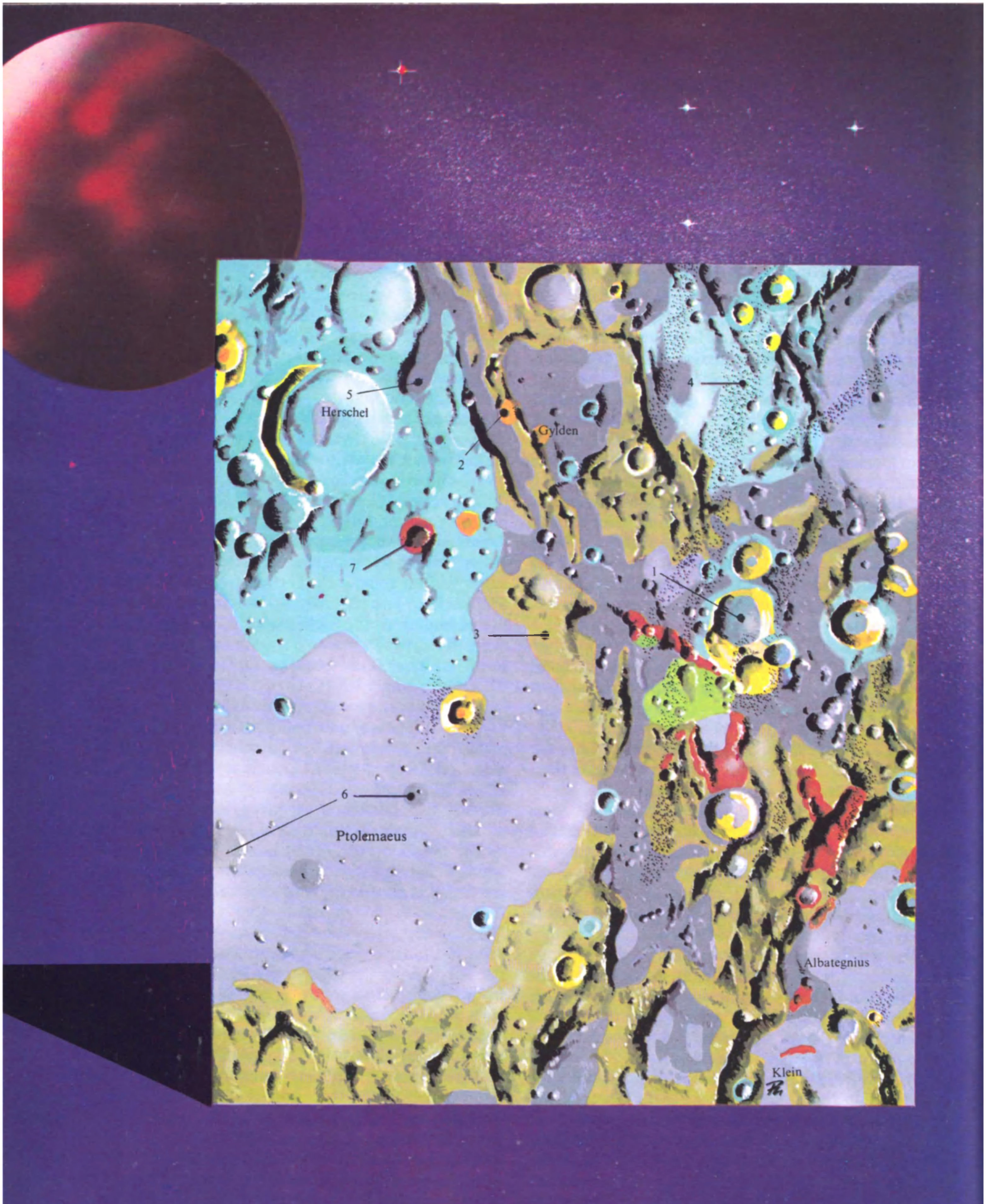
The Sun slowly journeys above the horizon, and it is 7.25 Earth-days after the lunar sunrise that the lunar midday arrives. Also, astronauts see our distant Earth in the sky. It is in the opposite phase to that of the Moon at the time.

When the Moon appears from the Earth to be completely dark and invisible near the Sun, our astronauts will see the Earth from the Moon as a bluish disk that is more than half-covered by white patches of clouds, so that it would be difficult to discern the continents on it. And the view, of course, will not resemble the "physical map of a hemisphere". They will only see dark oceans and light continents, there being no details as the colour contrasts will be smeared out by the mist of the terrestrial atmosphere. And the Earth's edges will be completely blurred by the atmosphere. The section "Travels To and On Hermes" will tell you more about what the Earth looks like from a distance. We will learn about how the activities of man on the Earth are seen from the Moon from the section "Is There Life on Earth?"

The surface area of the globe is 14 times that of the Moon, and as the Earth, with its clouds, is a better reflector than the Moon, the full Earth illuminates the night side of its satellite 40 times brighter than a full Moon illuminates us. Therefore, around new moon, during which the Moon appears as a narrow crescent, we can also perceive its dark (night) side illuminated by the reflected light of the Earth. This glow of the Moon is called ashen light. Against the ashen background we can easily see the contours of lunar maria.

Of especial interest are the observations of the Earth from the lunar poles and the other places seen from the Earth along the edge of the Moon's disk. Viewed from these places the Earth goes through some complex, though small, movements near the horizon because the Moon moves in a nonuniform way.

During new earth, the astronauts see the Earth either above or below the Sun as a light ring. The ring is caused by the atmosphere being illuminated by the Sun from behind. If the Earth happens to be just in front of the Sun, the astronauts will observe a total solar eclipse whose full phase will last 1.5 hours. The same event is perceived from the Earth as a lunar



5
Herschel
Gylden
2
4
7
1
3
6
Ptolemaeus
Albategnius
Klein

eclipse. During terrestrial solar eclipses there is vertually nothing to observe from the Moon – just a small dark spot with a half-shadow running over the Earth's surface, nothing to write home about really.

For the astronauts on the Moon the Sun makes a complete circle over the starry heaven relative to the horizon. As for the Earth, it remains in the same place somewhere above the horizon, depending of the observation point.

Some authors believe that astronauts exploring the Moon's surface should follow the terminator (the boundary between day and night), since in so doing the temperature will be the most favourable for them. At the equator the required speed would be 15.4 kilometres per hour and at the poles it would be possible to stay under these conditions even longer as sunlight always grazes the surface there.

The Soviet lunar probes have found that the Moon has no magnetic field like the Earth's. The terrestrial magnetic field conditions, as has been mentioned earlier, the presence of radiation belts around the Earth. The Moon has no such belt. It is only at certain locations that some magnetic fields are observed that are 1,000 times weaker than on Earth.

It is supposed that the magnetic field of the Earth is caused by electric currents within its liquid core. If this is really so, then the absence of a magnetic field on the Moon indicates that its bowels differ markedly from the Earth's core.

Since there is no magnetic field on the Moon a compass would be useless and astronauts can only take bearings from the positions of the Earth and the stars on the sky. To be able to tell the time and determine the coordinates of their location on the Moon, the astronauts will have to use some special methods, because terrestrial techniques will not work, for the Moon's axis is differently oriented in space, and its period and revolution around the Sun differ significantly from the Earth's. Some of these techniques have already been developed and they will continue to be developed.

Lunar geology at Ptolemaeus:
1 – a meteorite crater with hard rock external walls and loose boulders inside; *2* – a volcanic "dome" with flat elevations and small crater at the peak; *3* – a region with many small craters (500-2,000 metres in diameter); *4* – a ray system due to projectiles from impact sites; *5* – depressions without walls obviously formed by volcanic eruptions; *6* – mascons (meteorite craters that were later filled with lava); *7* – volcano crater.

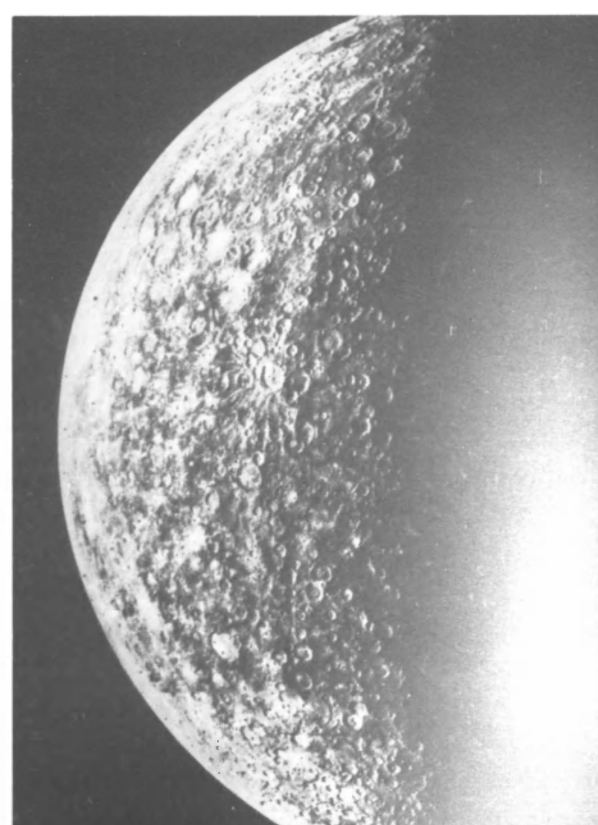
Two More Earth's Moons, but Made of Dust

The assumption that the whole surface of the Moon is covered by a thick layer of dust has not been confirmed, but as if by way of compensation, a Polish astronomer, Kordylewski, has found that the Earth has two more satellites, two more moons, and that these moons consist wholly of dust.

These satellites can be seen by the naked eye only from mountains and only when the air is especially clear, when they are suitably positioned above the horizon, and when the "real" Moon is above the horizon.

The discovery was made by Kordylewski as early as 1956, but during the ensuing 10 years only some people were lucky enough to see these satellites of the

Mercury from 950,000 kilometres (*Mariner 10* in March 1974).



Earth as exceedingly weak glowing spots that subtended an angle of several degrees. They travel among stars along the same path and with the same speed as the Moon, but one spot is always 60° ahead and the other 60° behind the Moon. This strongly suggests that both satellites lie at the same distance from the Earth as the Moon and form equilateral triangles with the Moon and the Earth. They lie at the so-called "libration points" predicted by Lagrange in the 18th century in his studies of the motion of three bodies of which two have large masses and the third has a small mass. Affected by the larger masses the third body will under arbitrary conditions quickly change its orbit. Its motion will be thus unstable. But if, according to Lagrange, the three bodies at the beginning of the motion are at the vertices of an equilateral triangle, their motion will be stable and their mutual arrangement will be retained. Small perturbations, for example the attraction of the others, more distant bodies, can however make the smaller body undergo slight oscillations about the vertex of the triangle, i. e. the so-called libration will occur.

The vertices of the two equilateral triangles just described are called the Lagrangian libration points. In the 20th century asteroids were discovered that move near the libration points in the system they form with Jupiter and the Sun.

It was not by chance that Kordylewski discovered the other Earth's satellites. He was looking for large meteorites near the libration points that could have got there from circumlunar space. Instead, he found the above-mentioned clouds and these could only be clouds of cosmic dust. They must be traps for dust stuck near the libration points in their disorderly movement in the solar system. The clouds reflect sunlight and must be very tenuous. They are somewhat larger in diameter than the Earth, their edges are not very sharp, and their mass is unknown but must be very small.

The Moon's Twin–Mercury

In terms of size, mass, proximity to the Sun, atmosphere, and mean density, all the planets can broadly be divided into two groups: Earth-type planets and Jupiter-type planets. The mean density of the Earth-type planets (and maybe of their satellites) grows as their diameter increases; the Moon has a density of 3.3 g/cm^3, Mars 3.9 g/cm^3, Mercury 5.4 g/cm^3, Venus 5.2 g/cm^3, and the Earth 5.5 g/cm^3. This is attributed to the relatively high abundance of heavy elements in the planets close to the Sun.

Mercury, the closest planet to the Sun, is difficult to study because it is for the most part lost in sunlight. Therefore, the early ideas of Mercury were mainly erroneous.

Mercury completes its journey around the Sun in 88 terrestrial days, it spins around the Sun "like crazy" from the "viewpoint" of the more distant planets, which ponderously revolve about our luminary.

When observed far from the Sun, Mercury appears through the telescope like our Moon in its first or last quarter, i. e. like a semicircle.

Observing some obscure spots on its surface, astronomers were led to the conclusion that Mercury, like a rabbit enchanted by a snake's stare, cannot turn away from the Sun and so faces it with one and the same side all the time. They even charted maps of the sunlit hemisphere of Mercury, though these were not given much credit.

Radar studies performed in 1965 in the USA showed unexpectedly that Mercury's period of axial rotation with respect to the stars was about 59 days. So Mercury turns with respect to the Sun in the forward direction, if only slowly. The planet has no hemisphere of eternal day and the hemisphere of eternal night, and its solar day is exactly twice its year.

But how is this to agree with the period derived from the earlier observation? Is there any contradiction here? None, the time spans between finding Mercurian spots at the same position on its disk for the same phases are satisfied not only by an 88-day period, but by other periods as well, including a 58.6-day period. This figure agrees with the radar evidence.

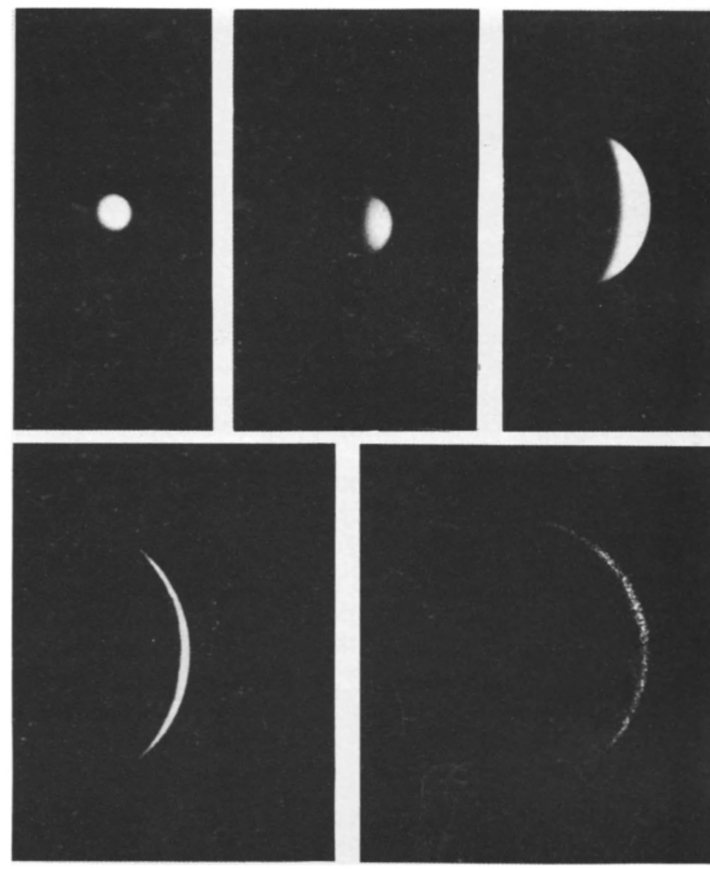
The Sun causes a tidal buldge on the planet always oriented towards it. If a planet rotates fast, this produces what is called tidal friction which slows down the rotation. If a planet's orbit is exactly circular, its rotation can completely stop with time. This was not the case with Mercury, however, since its orbit has a large eccentricity. These calculations based on the theory of gravitation, or rather on the theory of tides, even gave a quantitative explanation of the newly discovered axial period of Mercury.

In the light of the findings that on Mars, like on the Moon, there are many ring mountains, it was expected

that there would be many craters on Mercury as well, because it is intermediate in size between Mars and the Moon. This was confirmed by close-up pictures taken in 1974 by the US spacecraft *Mariner 10*, which flew past Mercury three times. Several thousand pictures of various scales were transmitted back that enabled much of the Mercurian surface to be charted in the same way as the Moon's. (Sometimes even in more detail than maps of some of the terrestrial areas.) As a matter of fact, the resolution of the pictures was as high as 100 metres, the length of large buildings on Earth. The whole of Mercury is as heavily cratered as the Moon. And there is no difference between the two planets in the sizes of craters, the steepness of their crater rims, both inner and outer, and the presence of smoother and darker bottoms of "maria" and "terrae". But unlike the Moon, Mercury is characterized by *escarps*, long ledges up to 3 kilometres high. Scientists associate the origin of such features with cracking and the dislodging of large areas of the Mercurian crust in earlier times as a result of the compression of the planet under internal transformations. Responsible for these surface features on Mercury are the same things as on the Moon, i. e. powerful volcanism and numerous falls of small and large meteorites in primaeval times.

Ring mountains and impact craters are formed in the absence of a sufficiently dense atmosphere capable of cushioning the impacts, for on this planet the gravitational pull is not strong enough for an Earth-type atmosphere to be retained. It was supposed that if Mercury had an atmosphere then it would be extremely tenuous. Neither direct observations nor Earth-based spectral studies have yielded any evidence for this atmosphere. It has only been found from a spacecraft that obtained spectra of some weakly glowing gases near Mercury's surface in ultraviolet light, which is caught by the Earth's atmosphere. The surveys of *Mariner 10* indicated that the concentration of gases at the Mercurian surface is many times higher than in interplanetary space. Abundances were estimated for oxygen and hydrogen, and also for the inert gases: argon, neon, xenon, and even helium. But the total pressure of these gases at the planet's surface is negligible, less than 1/500,000,000,000 of the pressure at the Earth's surface.

Mercury is 2.5 times closer to the Sun than the Earth, and thus receives 6 times more solar heat. Without an atmosphere of note and with its slow rotation it must be like a blast furnace during the day and at night it must cool down quickly for a long period. When the Sun is at zenith the Mercurian surface heats up to 290-420°C depending on the periodically varying distance from the Sun (owing to the large eccentricity of Mercury's orbit), and at the mean solar distance the temperature is 345°C. Hot like hell! Even when you roast something on a frying pan such a temperature is never reached, or everything would burn. When the planet is at its closest to the Sun, midday ponds of molten lead could appear at Mercury's equator.



The phases of Venus. The new Venus (between the Earth and the Sun): the crescent flickers by the light scattered in the dense atmosphere of Venus. In its later phases the planet becomes ever larger and fuller.

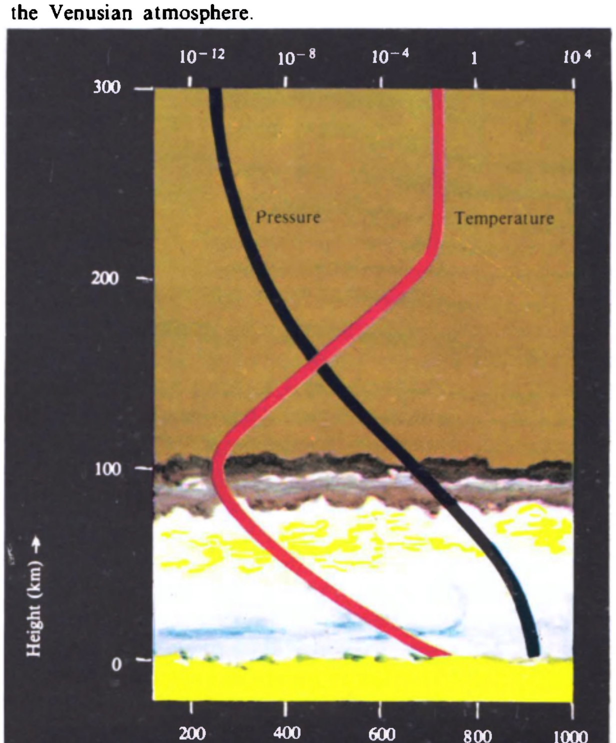
During the night, however, it is mightily cold. *Mariner 10* indicated that on the dark side the temperature dropped down to −173°C.

Mysteries of Beautiful Venus

Our neighbour Venus, only slightly the Earth's inferior in size and mass, is seen by the naked eye as a beautiful star. It bathes in sunlight of sunset and sunrise, and in ancient times it was taken for two different luminaries, called Vesper and Lucifer. Like Mercury, it changes its phases, but it wanders farther from the Sun, and so is more convenient to observe. When it is seen as a wide crescent from Earth, then it is at its brightest, and appears as a bright spot against the blue sky even during the day.

Venusian phases were discovered in 1610 by Galileo. He was not certain that his observations were correct and therefore was reluctant to report them openly, but at the same time he did not want to be deprived of priority. He therefore only published a Latin phrase that meant: "These immature things are already being handled by me, but in vain," and he added to this two letters that he could not use when he enciphered his

The variation of pressure (atm.) and temperature (K) with height in the Venusian atmosphere.



discovery. Galileo was asked on November 5 of that year whether Mercury and Venus have phases, as might be expected because, according to Copernicus, their orbits lay within the Earth's orbit. Galileo was very careful to answer that he had not yet studied all that concerned the heavens and that, being unwell, he had not been observing having to lie in bed instead. It was only after he conclusively established that his discovery was correct that he deciphered his anagram. Rearranging the letters it was possible to obtain the phrase (without extraneous letters) "The mother of love imitates the aspects of Cynthia." The mother of love was the goddess Venus, and Cynthia was one of the ancient names of the Moon.

Venus guards its secrets jealously. When it is close to us so that its narrow crescent can be seen even in strong binoculars, most of the hemisphere facing us is dark. And when it is illuminated almost fully, it is six times farther away and is lost in sunshine at that. Its atmosphere – the first thing to be discovered on it – contains a dense layer of clouds.

The white clouds of Venus are good reflectors of sunlight, which makes it the best luminary in the skies (besides the Sun and Moon). Venus exhibits no permanent spots and therefore neither visual nor photographic observations could be used to determine its period of axial rotation with certainty. Spectral analysis, too, gave no indication of noticeable rotation and so many observers decided in the long run that it, just like Mercury, always faces the Sun with one and the same side. Like Mercury... but radar explorations have shown that Mercury exposes all sections of its surface to the Sun. Radar studies of Venus have also revealed something unexpected. Both in the USSR and in the USA radioastronomers found that the sidereal rotation period of Venus was 243 terrestrial days, which almost equals its revolution period (225 days). But Venus rotates in the opposite direction, unlike any other planet, from Mercury to Saturn. Therefore, its solar day lasts 117 terrestrial days. Venus's inclination is very small, and so it has no seasons.

But let us return to the atmosphere with the clouds shrouding the planet. Spectral analysis has long since indicated that the Venusian atmosphere is heavy on carbon dioxide. Thereafter the data stopped coming. It is only in the 1960s, when improved astronomical instruments and spectroscopic techniques came in, that it became possible to find in the Venusian atmo-

sphere traces of other gases: a slight amount of water vapour and poisonous carbon monoxide, as well as hydrogen chloride and hydrogen fluoride. The last two are only encountered on Earth in volcano craters and chemical plants where they are used to produce strong acids. Unfortunately, Earth-based observations do not allow other gases to be found on Venus with certainty, even if they were the major constituents of the atmosphere. It is supposed that there is much nitrogen on that planet, just as on Earth. But testing this hypothesis awaits appropriate planetary probes.

Besides the chemical composition of an atmosphere it is important to know the temperature and pressure on the planet's surface. Earth-based observations of the thermal radiation of Venus in the infrared showed that the temperature is about −35°C both at night and by day. On the face of it, the result is shocking. Why is it so cold by day? It is known that Venus is 1.5 times closer to the Sun than the Earth. The fact is that the infrared radiation comes to us from the outer lining of the cloud "shroud" of Venus – it is their temperature that we measure from Earth. It is no wonder therefore because the temperature of the Earth's atmosphere, too, at high altitudes is low and varies but slightly from day to night and from the poles to the equator. These findings for Venus were substantiated in 1962 by the US probe *Mariner 2*, which passed the planet at a distance of 40,000 kilometres and beamed back some fascinating data, although not as much as had been expected.

At the centre of the planet's disk the temperature above the clouds was found to be −34°C, and at the planet's edges it reached −51°C, and sometimes even −57°C, thus suggesting higher temperatures inside the atmosphere. However, it was predicted from our knowledge of the physics of gases and planetary atmospheres that the temperature of Venus's lower atmosphere would be higher. But what is the thickness of the Venusian atmosphere and what are conditions like near the planet's surface?

It has long been obvious that the planet's atmosphere is denser than Earth's. It follows from the fact that when it is seen as a narrow crescent, the length of the crescent is much more than 180° along the limb, as is the case with the Moon. Under certain conditions (when for the Earth-based observer Venus almost lies against the background of the Sun) the ends of the crescent extend so that they close to produce a thin glowing ring. This is due to refraction (in Venus's atmosphere) and the high altitudes of the clouds scattering sunlight. Following different lines of reasoning scientists tried to work out the height reached by the clouds in the Venusian atmosphere. Some estimated it to be 30 kilometres, and others up to 100 kilometres. These two estimates give markedly different predictions of Venus's atmospheric pressure.

The continuous cloudiness of Venus, the Sun's neighbour, suggested that the conditions at the Venusian surface are analogous to what Earth must have been like during the Carboniferous Period in geological history. At the time, under the cover of clouds and in a damp atmosphere, rich in carbon dioxide, climatic conditions were favourable for the growth of exuberant vegetation. A hard blow against this conception has been delivered in recent years by radioastronomical studies of Venus. Repeated probings have shown that the conditions on this beautiful planet are far from beautiful for life. It is unbelievably hot there. From the planet's radio emission that can pass through the clouds and reach us from Venusian surface, the temperature was estimated to be 380°C. From millimetre waves, the temperature is lower (100°C), but this radiation comes from cooler parts of the planet's atmosphere.

The extremely high temperature, it is believed, is caused by the greenhouse effect: Venus's atmosphere is more transparent to optical than infrared light, and so it lets in sunlight that heats the planet's surface and atmosphere, but reflects back the heat radiated by the surface rocks.

One authority believes that part of Venusian surface in the hottest areas may be covered by molten metals such as lead (fusion point 327°), and maybe even zinc (419°C) and tin (231°C), which is less abundant there than on Earth, and may also be covered with molten chemical compounds belonging to the class of low-melting carbonates. We do not know if there are such lakes on Venus, but it turned out that on day and night sides, and even at the poles and equator, the temperature is virtually the same, the difference being as low as several degrees.

Under such conditions a landing of astronauts on the planet is impossible. Probes are "tougher" than man, but they too require protection from the excessive heat.

On October 18, 1967 the Soviet spacecraft *Venera 4*


1
2
3
СССР

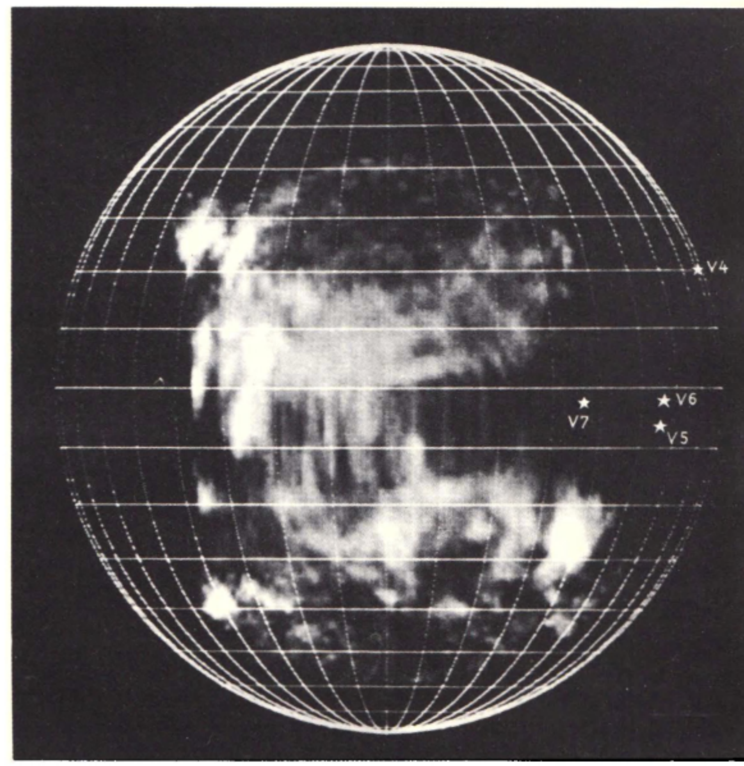
took an instrument capsule to the planet, which parachuted deep into the atmosphere for the first time (the day before *Mariner 5* passed at 4,000 kilometres from Venusian surface and reported back scientific data). The capsule made direct measurements inside the planet's atmosphere. The data obtained were supplemented and refined using improved apparatus in subsequent experiments. In 1969 *Venera 5* and *Venera 6* parachuted through the Venusian atmosphere on the night side. Radio signals stopped when the temperature rose to 325°C and the pressure to 27 atmospheres. One of the major results of these experiments was the discovery that carbon dioxide in the planet's atmosphere is its major constituent, and that its nitrogen content by volume is not higher than 2 per cent, that of oxygen is not higher than 0.1 per cent, whereas water vapour near the cloud layer accounts for only a few tenths of a per cent.

In 1970 *Venera 7* succeeded in softlanding on the surface of Venus, also on its night side, and sent data from the surface (obviously, solid). At the landing sight the temperature was found to be 470°C and the pressure 90 atmospheres with the possible error of $\pm$ 15 atmospheres. The atmosphere was 60 times denser than on Earth. Unlike its sister probes, *Venera 8* landed in 1972 onto the day side of Venus and reported the temperature to be 470 $\pm$ 20°C and pressure 93 $\pm$ 1.5 atmospheres. The illuminance meter indicated that, due to dense clouds, there was dusk by day on the surface of the "morning star", but between the clouds and the surface there were some transparent layers in the atmosphere.

The first TV panoramas of the Venusian surface were transmitted in October 1975 by instrument capsules from the Soviet space probes *Venera 9* and *Venera 10*. In December 1978, Venus was attacked by a swarm of scientific probes among which were four capsules that separated from the American Venus satellite *Pioneer-Venus* and two Soviet spacecraft *Venera 11* and *Venera 12* equipped with improved apparata to study the Venusian atmosphere. The probes enabled the amount of nitrogen to be ascer-

Method of soft-landing an instrument capsule from the *Venera* spacecraft:

1 radar altitude measuring; *2* start of communications session with Earth; *3* braking, and then main parachutes are pulled out. The capsules of *Venera 9* and *Venera 10* used radically new soft-landing systems. The parachutes were discarded at 50 kilometres and the aerodynamically retarded capsules safely landed. The impact energy was absorbed by thin-walled envelopes.



Radar picture of Venus.

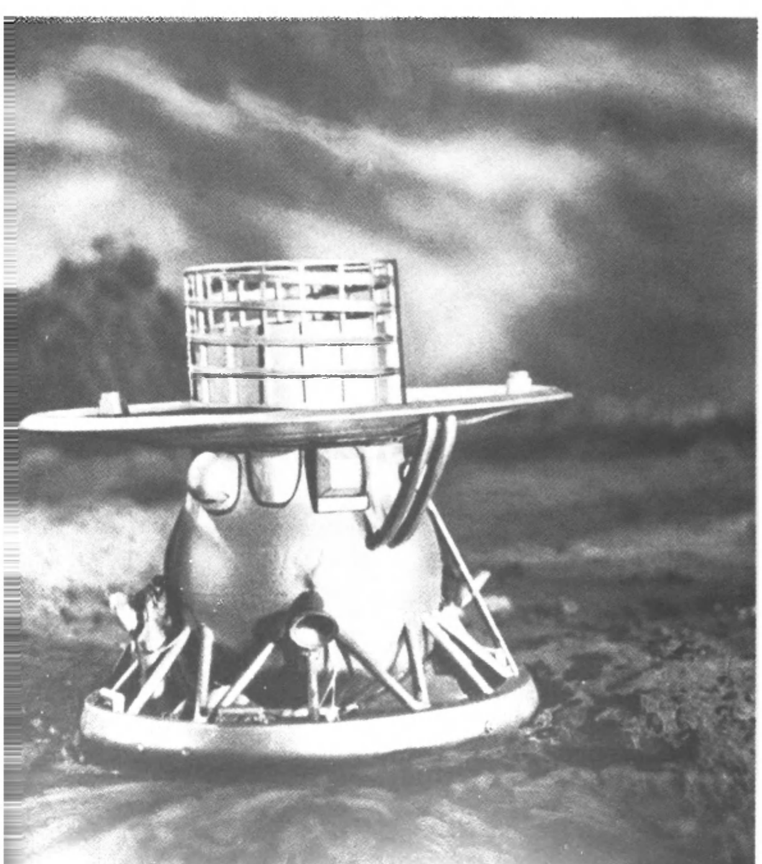
tained (about 4 per cent), to measure the content of carbon monoxide (about 0.01 per cent) and to find inert gases (argon, neon, krypton in trace quantities), and also sulphur compounds.

A remarkable achievement was the isotopic study of the gases. The isotopic-abundance ratio for oxygen in the Venusian atmosphere was found to be nearly the same as in the Earth's atmosphere, but with argon the ratio appeared to be markedly at variance. These findings provide a clue to the planet's origin. It is well known that the isotopes argon-36 and argon-38 are not born on the planets, hence they were part of the "parent" matter involved in the formation of the planet. But argon-40 is constantly produced on the planet as the potassium-40 in the rocks decays. The amount of primary argon isotopes on Venus has been found to be approximately equal to that of the radiogenic argon-40, whereas in the Earth's atmosphere the ratio is 1 : 300. The reasons behind these differences have not yet been understood.

We have deliberately given here a brief account of

The capsule of a Soviet spacecraft parachutes to softland.

The lander of *Venera 9* and *Venera 10*.

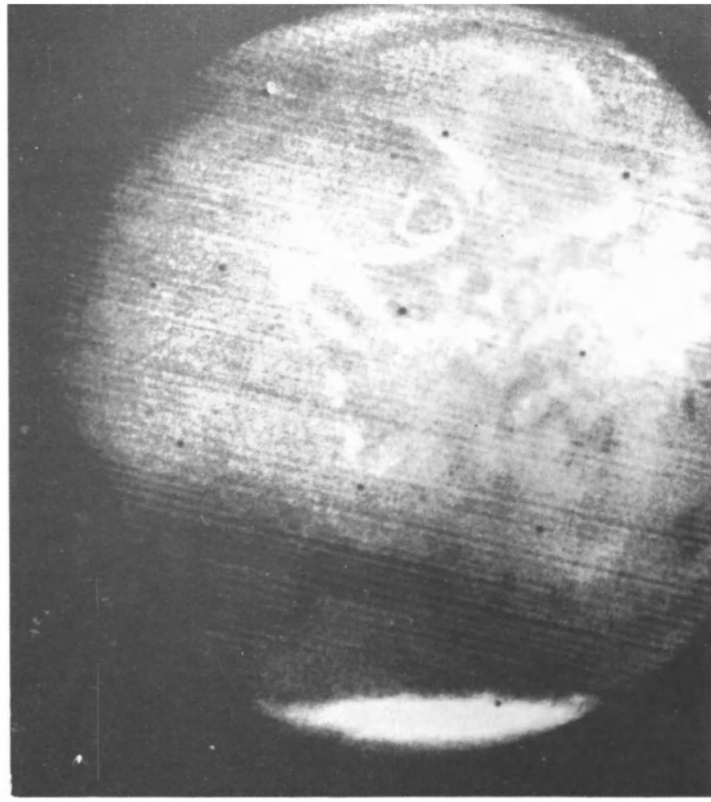


the history of Venusian studies so that the reader could appreciate the time and effort required to explore the planets and the way in which aeronautics supports the inferences drawn from Earth-based observations.

Now let us return to Venus's clouds. They float very high over the planet's surface, their lower boundary being at an altitude of 47-48 kilometres. Beneath the clouds the dense atmosphere is more transparent, but it still absorbs light significantly, so that no more than 10 per cent of the incident radiation reaches the bottom of the gas ocean of Venus, even when the Sun is at its zenith. Especially absorbed are blue and green rays and so the sky looks wine-coloured from the Venusian surface. This absorption is ascribed to the presence of gaseous sulphur formed in the decomposition of chemical compounds in the hottest strata of the atmosphere up to 22 kilometres from the surface. The composition of Venus's clouds has not yet been established though now it seems unlikely that they will consist of water droplets or ice crystals. In recent years the hypothesis has gained recognition that these are droplets of an aqueous solution of sulphuric acid. This assumption accounts for the amazingly small amount of water vapour in the atmosphere and agrees with optical observations.

Ultraviolet pictures of Venusian clouds made as early as the 1920s showed some large darkish spots. Comparing pictures obtained at different times, astronomers suspected retrograde rotation (in an unusual direction) of the planet, which according to the best estimates corresponded to a 4-day period. The conclusion was borne out by the data supplied by *Mariner 10* in 1974, which approached Venus and sent back ultraviolet pictures.

Radar surveys of the Venusian surface have enabled maps to be charted. Later American scientists used the radar on Venus's artificial satellite *Pioneer-Venus* to reveal mountains in the northern hemisphere of Venus with an elevation of 5-10 kilometres with respect to the average level of the surface. In the southern hemisphere they found an elongated feature similar to *rift valleys* along faults in the planet's crust consisting of long narrow collapse basins and narrow uplands along these faults, the height differences being up to 7 kilometres. It came more as a surprise to find craters on Venus that are from 50 to 250 kilometres across. Some of them may be of volcanic origin, others of impact



Mars with its dark and light parts and "snow-covered" southern polar cap (*Mariner 7*).

origin. Of course, there must be plenty of smaller craters, but it is unlikely that there will be meteorite craters less than one kilometre. The dense atmosphere of Venus is impenetrable for small meteorites, but at the same time it cannot stop larger bodies (one kilometre or more across) producing craters with a diameter of several kilometres. Venusian craters are not deep: even the largest have a depth of only 0.5 kilometre. Obviously, the flattened surface is a result of weathering.

So the beautiful Venus is gradually uncovering her mysteries.

Mars from Far and Near

Oh Mars! A fascinating planet!

The distance from the Earth to our celestial neighbour varies widely. It is the closest at a so-called opposition, when both planets, travelling along their respective orbit, approach each other on the same side of the Sun. Oppositions occur once every two years. But the separation at oppositions also depends on where along their orbits they are when they approach each other, since the orbits of both planets, especially of Mars, are noticeably elliptical. So the smallest separation varies from 100,000,000 to 55,000,000 kilometres. In the latter case the opposition is called favourable and recurs at intervals of 15 or 17 years. The next favourable opposition is to occur in 1988. Unfortunately, at these oppositions Mars does not appear high above the horizon for the middle latitudes of the northern hemisphere where most of the observatories are located. It thus has to be observed through a thick layer of air – the perennial enemy of astronomical observers.

Mars is 1.5 times farther away from the Sun than the Earth, and it receives about half as much heat from it than the Earth. It completes its orbit about the Sun once in 687 Earth-days (or 668 Mars-days) and spins on its axis with a period only slightly longer (by 37 minutes) than a terrestrial day. Mars's axis is inclined to its orbit almost like the Earth's, so that it has its seasons. Lastly, the Martian diameter is half that of the Earth and its gravitational pull is half that of the Earth. Therefore, it is much simpler to launch spacecraft from the Martian surface.

Owing to air perturbation we cannot from Earth resolve details on the Martian surface less than 0.5-1″, or less than 150-300 kilometres across even through a strong telescope. Moreover, Mars has an atmosphere of its own, which, although tenuous, causes a mist above its surface, especially on its edges. Therefore, it is hardly possible to see much on Mars, and only patient observations during the best nights of the year provide some overall ideas of the surface of this enigmatic planet.

What then can you observe directly on Mars? In the first place we note that most of its surface is yellow-reddish, the colour of blood, which is why the ancients devoted the planet to Mars, the god of war. These areas were earlier believed to be deserts, smooth and plain. Another distinctive feature is a polar cap on one of Mars's poles.

The white polar caps are a thin layer of snow often covered by fog, according to the 1909 hypothesis of the Russian astronomer Gavriil Tikhov (1875-1960). He was the first to observe and photograph Mars through different light filters. It was suspected that the caps are seasonal covers, because when it was winter in one hemisphere the cap was large, and towards the summer it fragmented at the edges and sometimes even vanished.

In recent years it has been found that besides "water" snow the Martian polar caps contain (and are perhaps dominated by) the snow of frozen carbon

dioxide, which under the conditions on the planet, solidifies at about $-124°C$.

It is more difficult to perceive dark spots on Mars because there is insufficient contrast. From the very beginning they were called maria, or seas, but, of course, they are not. The Sun's reflection has never been observed in them. The light-orange regions between the dark spots were called terrae. Accurate measurements indicated that the maria have almost the same colour as the terrae, reddish.

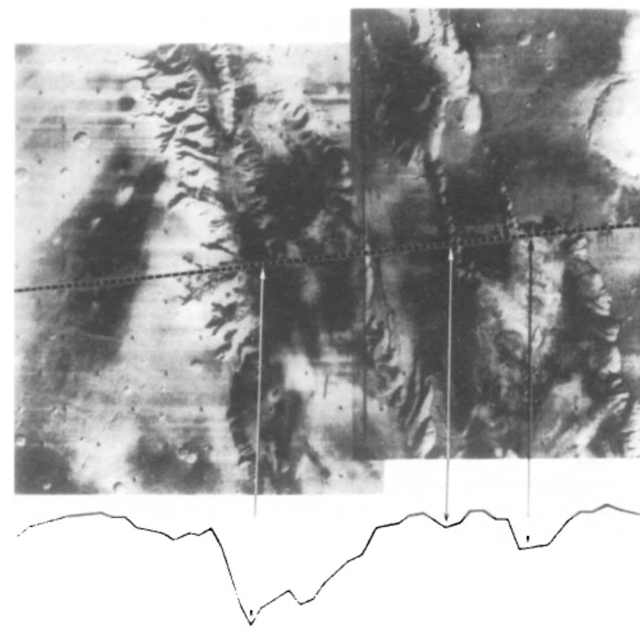
The orange continents were believed to be smooth deserts, and the mountains, it was held, were only where there were scanty white spots – perhaps snow in the mountains.

The major spots are recognizable in all the maps of Mars based on the photographs and best drawings of the planets made in astronomical observatories during the last 90 years by about 200 observers. Individual drawings showed details on Mars to within 5° (600 kilometres) in longitude and latitude. But the maps obtained from Earth-based observations do not show details of relief but only the dark and light regions no less than 150 kilometres long, sufficiently large to be seen in the telescope.

The Earth-based observations of fine features on Mars have been performed at the limit of astronomical optics and human vision, using the modern parlance at small signal-to-noise ratios. Under these conditions various spurious effects such as optical illusions due to physiological behaviour of the eye, begin to manifest themselves.

Giovanni Schiaparelli (1835-1910) of Italy saw in 1877 a network of thin lines on Mars, which he called canals. The lines crossed not only the orange terrae connecting the dark spots, but the seas too. Percival Lowell (1855-1916) of the USA believed that the canals were artificial structures built by hypothetical Martians, evidence of their highly developed technology. He drew a geometrically ideal network of hundreds of canals. But Antoniadi and co-workers found that the canals were not thin lines, but were hundreds of kilometres wide. Through large telescopes the canals appeared as a collection of irregular spots, rather than lines. Then it was supposed that the canals were patches of vegetation lining narrow water courses, which were not visible in the telescope.

To answer the many Martian puzzles a more detailed picture of the planet's surface was necessary.



A Mars-scape – the Grand Canyon.

Our ideas of the Martian surface began to change after the 1965 fly-by of the US spacecraft *Mariner 4*. It took eight months to approach Mars and from 12,000 kilometres it took and transmitted back to Earth 20 yet rather unsatisfactory pictures of the planet's surface. They covered only 1 per cent of the Martian surface along an arch passing through both dark and light regions, but much to the surprise of the astronomers, it was found that both light and dark surfaces were heavily cratered, just like the Moon. The craters on Mercury were found nine years later. In 1969 more pictures were obtained by the probes *Mariner 6* and *Mariner 7*.

Subsequent years saw the launches of several Martian probes: the American *Mariner 9* and *Viking 1* and *2*, and the Soviet *Mars 2, 3, 4, 5, 6,* and *7*. Nearly all of them became Mars's satellites and enabled the planet's surface to be observed for a long time from a fairly close range. The Mars landers *Viking 1* and *2* performed a large programme of surface studies directly on the Martian surface at two widely separated points. These detailed studies on the surface were combined with the essentially global overview of the entire sur-

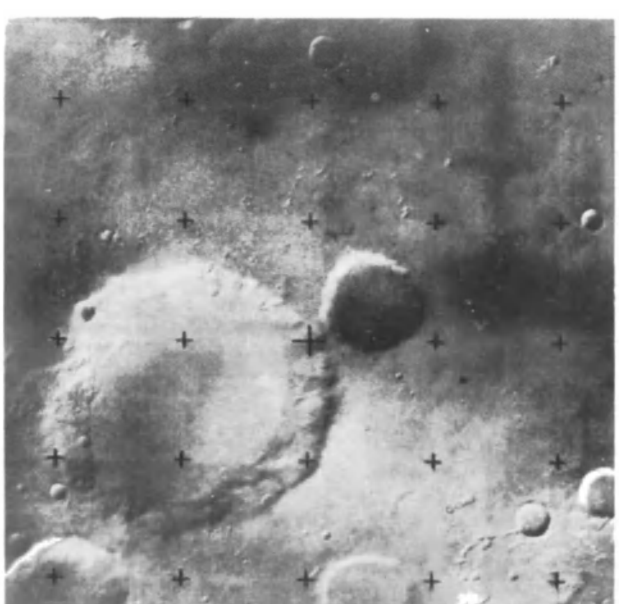
face of the planet from the satellites. Satellite-based observations at times were hindered by dust storms on Mars which enveloped its surface and saturated its atmosphere with fine dust. Sometimes the Martian atmosphere remained so dusted that little was perceived through it for months. The dust storms were a telltale sign of an atmosphere on Mars and had been observed long before spacecraft observations and landings. It is common knowledge that for dust to be transported winds are required. The transport of sand by seasonal winds can in principle explain the variation of the outlines of the Martian dark spots, long observed from the Earth. The yellow clouds of dust agreed with the notion that the yellow "plains" on Mars are highlands and sand deserts, and the dark "plains" are lowlands that perhaps have water and vegetation. It is now established that there are height differences on Mars more than 20 kilometres (like what we have on Earth). But the first measurements of height difference for large regions indicated that highlands and lowlands do not coincide with the orange "deserts" and "seas", i. e. with the dark spots. In no case do the boundaries of the dark spots coincide with the boundaries of the surface features found in the pictures of Mars obtained by the American and Soviet spacecraft.

Of the numerous craters found in the pictures more than one hundred are more than 120 kilometres across. Smaller craters are more numerous but not as numerous as they are on the Moon or Mercury. Their slopes are more gentle, being often about 10°. Central hills and craters on rims are rare occasions. This is all the result of erosion due to the action of winds, impacts of small meteorites and sand storms, and in the past, perhaps, of water.

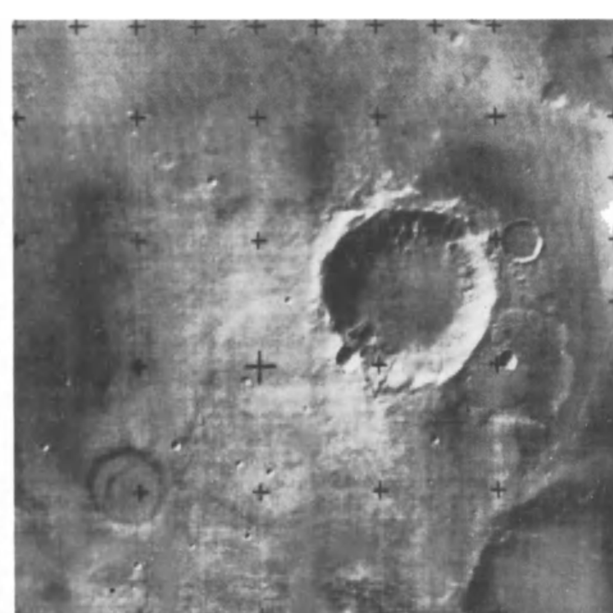
The record-breaker among the likes is the Grand Canyon *rift valley*, later renamed Valles Marineris, the Valley of Mariners. It is several thousand kilometres long, 100 kilometres wide, and several kilometres deep. It has no counterparts on the Earth or Moon. Another unique feature is Nix Olympica, the Snows of Olympus. This is a large volcanic cone similar to the ones on the Hawaiian Islands, but larger by far. Its base is 500 kilometres across and on its summit, like in Hawaii, it seems there is a lava lake, a caldera. There are several other shields of this type in this region, although smaller ones.

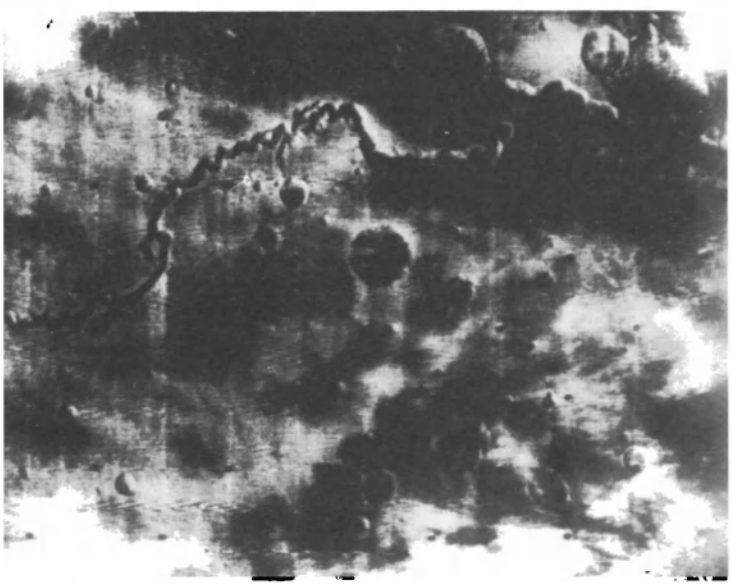
These newly found topographic features do not dis-

A flat crater on Mars (*Mars 5*).



A Martian crater with a small crater in the wall (*Mars 5*).





An ancient river-bed on Mars.

agree with our earlier ideas of the natural conditions on Mars, they only supplement and refine these ideas. But nobody expected to find there some meandering valleys and tributaries, which have to be recognized as the *courses of rivers that at one time flowed on Mars* (on a planet where it is now difficult to find water vapour!). It is quite obvious that Mars was once so rich in water that it could flow over the planet's surface. Why was it so and why is it not so now? The answer has to await further study. An interesting hypothesis was put forward by Davydov. He postulated that there are still frozen water reservoirs on Mars covered with sand. But where? For example, under the smooth surface of some regions in a lowland in the moderate latitudes of the planet's southern hemisphere.

The weather conditions on Mars are as yet insufficiently studied, but in general it is, brr!, very cold there. Bolometer and other studies of the thermal, or infrared, emission of the planet have yielded the following results. Even at the equator the temperature varies periodically owing to significant variations of the Mars-to-Sun distance. During the coldest period, at aphelion, the temperature of the Martian surface at the equator is never more than 0°C during the day. In the warmest season it reaches 15-20°C, but by sunset it drops to freezing temperatures and at night it falls to −100°C or lower. The temperature of the dark markings is somewhat higher than that of light regions (by 10°C) since they absorb sunlight better. The mean daily temperature there is like in the permafrost regions on Earth: −25°C (for the Earth as a whole the mean yearly temperature is +15°C). During the summer the daytime temperature on the Martian slopes that face the Sun is near zero even in the polar regions, where in winter it drops down to −130°C. Radio emission studies show that for the planet as a whole the temperature is about −70°C. Obviously, this low temperature refers to the subsurface layers since it is nearly the same whether Mars is illuminated by the Sun or not.

Geophysics teaches that the temperature of Martian soil must increase with depth in approximately the same manner as on Earth where it increases at a rate of about 30°C per each kilometre of depth.

The temperature on Mars is slightly influenced by

The volcanic caldera Olympus Mons.

its atmosphere. Mars's atmosphere, like Venus's, is mostly composed of carbon dioxide. It was found from spectral analysis that the planet's atmosphere contains about 0.1 per cent of oxygen and carbon monoxide. Since terrestrial oxygen is thought of as the result of the accumulation of millions of years of vegetation activity, the paucity of oxygen on Mars is evidence against the hypothesis that in the past there was plenty of vegetation there.

In 1963 the content of water vapour in the Martian atmosphere was measured for the first time. If converted to liquid, the amount of vapour would yield a layer of only 10-20 micrometres. In the Antarctic in frosty weather we can find about as much vapour in the air.

Chemical analysis of gas samples from the landers *Viking 1* and *2* allowed more accurate data to be obtained on the atmosphere's composition: 2.7 per cent nitrogen, and 1.6 per cent argon. At the surface the pressure is about 6 millibar (1 millibar equals 0.75 mm Hg). In the terrestrial atmosphere such pressures are only encountered at altitudes of approximately 50 kilometres. In Martian mountains and lowlands the pressure is different and it undergoes seasonal and diurnal variations. The same refers to temperature.

In the coldest parts of Mars, apart from hoar of water ice crystals, hoar of carbon dioxide also occurs. It has a tendency to condense mostly in lower regions, not on the summits. Both the summits and valleys have nearly the same temperature in the mornings. But for carbon dioxide to condense on the summits where the pressure is lower, much lower temperatures are required than at ionosphere base. It so happens that instead of white "caps" Martian mountains must wear white "collars".

The landers gave the first opportunity to see the colour of a Martian sky. During the day it is pink. This is associated with the light scattering by aerosoles, i. e. fine dust particles suspended in the atmosphere.

Mars orbiters have provided evidence for the ionosphere in the upper atmosphere of the planet. Just as on the Earth, this layer contains many ions (hence its name) and, strictly speaking, consists of several layers, the main layer lying at an altitude of 120 kilometres. By day the ion density is as in the Earth's E-layer by night (10^5electrons per cubic centimetre), and during the night it is about 20 times smaller.

But Mars has no radiation belt similar to that around the Earth. This is accounted for too small a magnetic field on Mars, as was found by *Mars 2* and *Mars 3*. The Martian magnetic field is 500 times or so weaker than the Earth's, but it is still there, perhaps as a result of the convection of matter deep inside Mars, which is only possible with molten liquid core in the planet.

The weak magnetic field, like a soap bubble in an air current, can be markedly deformed under the action of gusts of "solar wind", when the planet is bombarded by streams of electrically charged particles ejected by the Sun. Therefore future visiters of Mars will not rely on a compass and will have to use more accurate navigational devices. Also, the moving about the heavily cratered Martian surface will be nearly as hard as over the Moon. It is well known however that on the Moon the Soviet moon cars and American lunar roving vehicles worked and moved successfully. It was reported that in the USA special robots have been developed to travel about the planets in a vacuum or a dusty atmosphere.

To be able to work out robots for exploration on Mars and life support systems for astronauts a good knowledge is required of the physical conditions that they will encounter there. Besides, an understanding of the physical conditions on Mars will enable us to be on firmer ground when discussing the issue of life on it.

Once More About Life on Mars

Fairly recently the first question acquaintances would ask half in jest when meeting an astronomer was, "Well, what's happening there on Mars?" So many novels, both bad and good, have been written about the inhabitants of Mars, the topic also intrigued the producers of fantastic films. So I am obliged to dwell on this theme too.

The question of whether or not there is life somewhere else in the solar system is a question of enormous importance for our outlook. The heated disputes about the presence of life on Mars and what it may be like have been an attraction for scientists in various branches and all those interested in the science of the Universe for about a century, especially after the discovery of "canals" on Mars, whose geometrical regularity and artificial origin as advocated by Lowell and others, have long been disproved. We have already

discussed the issue earlier. Now this viewpoint can only be held by ignorant people or imaginative science fiction writers.

The issue of life basically reduces to the following three questions: (1) Could some sort of life have originated on Mars? (2) Can it exist there now? (3) Are there any signs of its existence?

The first two questions, in their scientific form, can only be answered by relying on the notion that, as on Earth, life is solely possible on protein basis, on hydrocarbon compounds. Whether some other kind of life is possible, on some other basis, is not known. There is no unified idea of how life came into being on Earth, and our conception of the conditions on Mars milliards of years ago is very hypothetical. So the situation does not warrant any positive statements, but in the majority of cases conclusions are negative. Under current conditions life is not possible on Mars. It has not yet been established whether some such possibility has occurred in the Martian past, although there are some promising indications. An interplanetary transfer of bacteria and spores is highly unlikely and would require special conditions. If it does occur then the organisms are bound to be destroyed by cosmic and X-rays in interplanetary space.

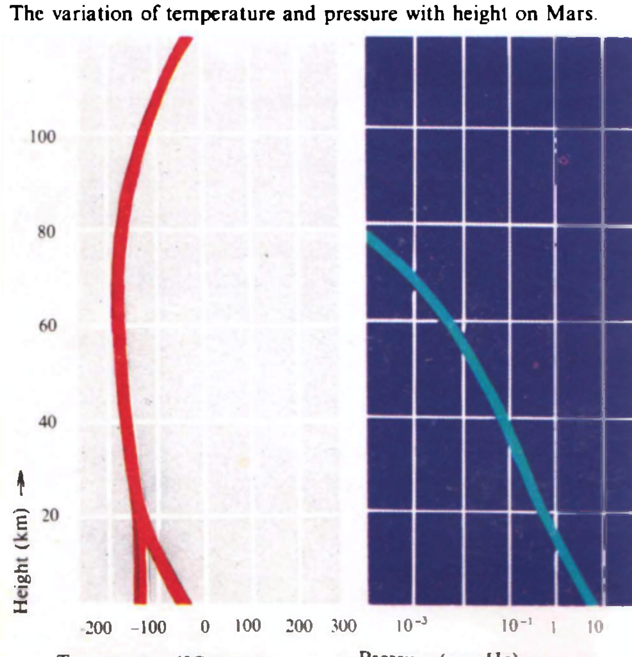
Nevertheless until fairly recently scientists believed that at the present time some life was possible on Mars and that there were even indications to this effect. Although the conditions on Mars are extremely inhospitable, some scientists refer to the unfathomable adaptation power of some life forms, specifically to low humidity and low temperature, as well as to its fluctuations. Admittedly, the most hardy are bacteria and lower plants.

The surest way to test the hypotheses about life on Mars is to carry out studies directly on its surface. The first attempt to search for extraterrestrial life was undertaken in 1976 by the American landers *Viking 1* and *Viking 2*. The Martian panoramas they sent back to Earth show a chaos of stones in a sand desert without any signs of life. Both stations were equipped with apparatus to search for some indications of metabolism of Martian microorganisms and to perform chemical analysis of rock samples. The samples of fine sand analysed by the robots contain (by weight) 44 per cent silicium dioxide, which is the major constituent of our sand, the balance being oxides of iron and other widely occurring metals – aluminium, magnesium, calcium – with the exception of potassium (its absence is not as yet understood) and, quite unexpectedly, a large amount of sulphur compounds. No microorganisms have yet been found on Mars. The samples studied contain no residues of organic matter, but the samples were taken at two sites only. So some optimists accept the possibility that maybe life will be found in some other, as yet unexplored, places on Mars.

A prominent Russian physicist, Nikolai Umov (1846-1915), indicated back in the last century that if a planet has some vegetation, the spectrum of reflected sunlight should display an absorption line of chlorophyll. Chlorophyll is a green pigment in plants that absorbs infrared light and it shows a wide line. The spectrum of the dark Martian markings exhibited no chlorophyll lines.

The advocates of life on Mars were alerted in 1956 when Sinton of the USA reported the discovery in the infrared spectrum of the Martian maria of three absorption lines similar to those observed with organic matter. But Mars scientists found in 1963 that the same spectral range also contain absorption lines of inorganic carbonates, e. g. limestone. Other research workers found in 1965 that the Sinton lines are also

The variation of temperature and pressure with height on Mars.



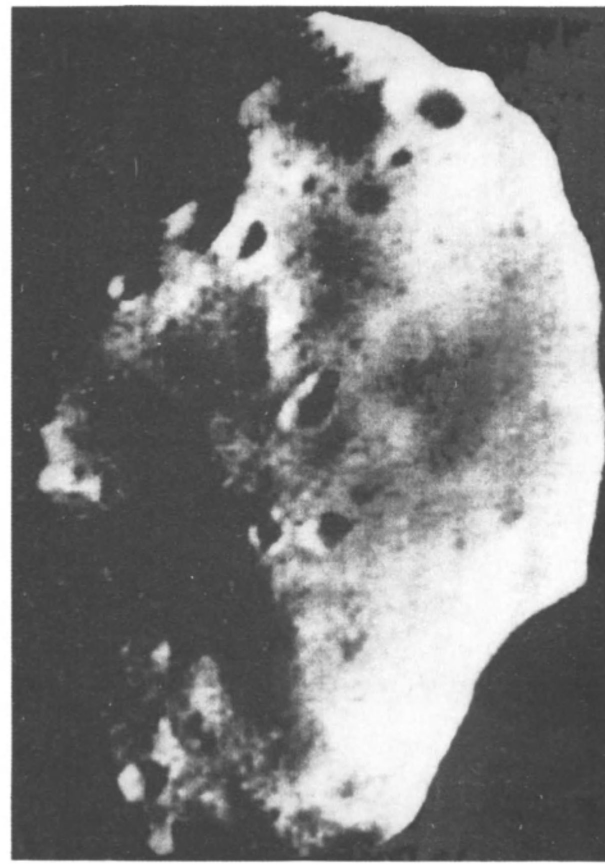
produced by heavy water, which is ordinary water with deuterium substituted for common hydrogen. An estimate of the amount of heavy water required to yield the Sinton lines coincided with the estimate we've given for the amount of water vapour in the Martian atmosphere. But this suggests that on Mars the amounts of hydrogen and deuterium must be the same, whereas on Earth hydrogen is 5,000 times more abundant. Attempts to account for the difference have not yet been convincing. Some authors believe that the Sinton lines are produced by HDO molecules (where D is deuterium) in the terrestrial atmosphere when it is rich in water vapour. Why then are these lines seen in the spectrum of maria and are absent in the spectrum of deserts on Mars? Later, however, Sinton reported that the maria and the deserts had been observed on different nights and thus under different conditions in the terrestrial atmosphere. The Sinton lines are in need of further studies.

It is hard to expect finding higher plants on Mars. Clearly, if plants are to be found there at all, they will only be some form of mosses and lichens. It is even more difficult to make assumptions about any form of animal life on Mars, especially intelligent life.

Are the Martian Satellites Artificial?

Two satellites of Mars were found in 1877, and after the companions of the "god of war" they were called Deimos and Phobos, which is the Greek for "terror" and "fear". But Terror is terribly small, and Fear is smaller still. The former is under 27 kilometres across and the latter 16 kilometres. They are comparable in size to the lesser planets, the asteroids. It may well be that they indeed were asteroids taken prisoner by Mars. Incidentally, they had been suspected twice without any reason. As we will see in the story of the Saturn's rings, Kepler erroneously thought that Galileo's coded phrase about his discovery of the planet Saturn meant Galileo had discovered Mars's two satellites. That was in the 17th century, and in the 18th century the English writer Swift made his fictitious astronomers of Laputa discover two satellites of Mars.

Deimos revolves very near its master planet, 23,500 kilometres from its surface, and Phobos is closer still, 5,400 kilometres from the surface. Therefore, during the Martian day Phobos rises twice above the horizon and passes through all its phases, the Moon-like fashion. In so doing, it rises in the west and sets in the east, its revolution period being only 7 hours 37 minutes. (The Earth's artificial satellites behave in much the same way.)



Phobos.

Some observers have found that Phobos's revolution period is decreasing by one-millionth of a second per day. Shklovsky concluded in 1960 that the drag of the Martian atmosphere is responsible and for the braking to be as observed the mass of Phobos must be extremely small. This suggests then that its average density is a thousand times smaller than that of water. That is only possible if, having a solid surface, Phobos is empty inside. But then it can only be artificial.

Did Shklovsky attach much importance to this conclusion or did he amuse himself at the sensation it produced in the press? In the light of the boom of the

artifical terrestrial satellites the idea was fashionable.

There are natural explanations for the motion of Mars's moons. Some scientists found that if the Martian crust is less tough than the Earth's, then the tides due to Phobos can slow it down in accordance with the observations. On the other hand, V. V. Radzievsky with co-workers showed that if the Martian satellites are substantially different from spheres, then the pressure of sunlight is also more than enough to accelerate Phobos and slow Deimos in accordance with the observations.

Consequently, Mars has no artificial satellite. Or has it? No! But we have almost forgotten that it does have artificial satellites, and many of them at that, although they were not made by the hypothetical inhabitants of Mars but by the dwellers of the Earth. Without being exposed to noticeable perturbations they will revolve about Mars for hundreds, or even hundreds of millions of years. True, they are small even in comparison with Phobos and Deimos, but Deimos and Phobos are themselves negligible as compared with Ganymede, a Jupiter's satellite that is larger than Mercury.

Finally, only ten years after the hullaballoo about the allegedly artificial satellites of Mars, this makeshift hypothesis was demolished by close-ups sent back by the real artificial satellites. The pictures positively showed that both satellites are absolutely irregular in shape, just like potatoes. They are both pieces of stone rubble. But the most important and striking thing was that both satellites appeared to be covered with craters. The largest of the craters on Deimos, with a diameter of about two kilometres, covers nearly a quarter of its "hemisphere".

To be sure, the volcanic origin of the craters here is absolutely out of the question, and so they have been produced by meteoritic impacts in the remote past. As for their smooth surface, Phobos and Deimos had it "processed" by impacts of small meteoritic bodies. Phobos also appeared to be covered by a system of deep parallel troughs up to 1,000 metres long and up to 20 metres deep.

In the last century hyperactive minds assumed that the light spots that sometimes appeared in certain places on the surface of Mars are light signals sent by Martians to us people. But these appear just to be mountains covered with frost. The first radio noise from space was also ascribed to Martian signals. I would say that if the Martians were in existence and tried to establish some sort of communication with the inhabitants of the Earth, they would have long given it up as a bad job and stopped signalling. If they had surpassed us in technology long ago, they must have been disappointed because even as recently as twenty years ago we could not have answered them. Was it really worth sending signals for millennia without receiving a reply?

There would be nothing threatening for the philosophical idea of life if it turned out that within the solar system there is only life on Earth. After all the solar system is not the whole Universe and in our (in our alone!) stellar system, called the Galaxy, there are more than one hundred milliard stars, and if only one in a thousand would appear to have an inhabited planet, this would yield one hundred million inhabited planets. Multiply this by the almost infinite number of galaxies in the Universe and you will arrive at an infinite number of inhabited planets. The Russian poet A. Kovalenkov expressed the hope thus:

Twinkling mysteriously in our eyepieces
Through the darkness of the cosmic depths sails
Our globe's orange neighbour,
The lord of fantasies and utopias,
Mars: by millions of miles of distance
Diminished, turned into a child's ball,
It sails, twinkling with the fading light
Of twilights unknown to us.
What is hidden in its silent deserts?
What kind of life that to avoid being crisply burnt
A network of canals – a system of strange lines
From pole to equator have woven.
Who breathes the thin air there
And, perhaps, has watched for how many ages
Our green planet
Where the Martian was invented by man?
Three-legged giants of frightening stories,
Iron mushrooms in the crimsom gloom...
Wells made them as a warning to people,
Frightening too bold minds.
No, not for this did the dream seek light
And tug at Tsiolkovsky's heart,
To send a rocket as a herald of war
There where life was waiting for salvation.
We cannot divine or fix the time
But it will come – this day and hour
When we shall meet those who from afar
Without losing hope, believed in us.

The view of the Earth from space (*Zond 7*, 1969).

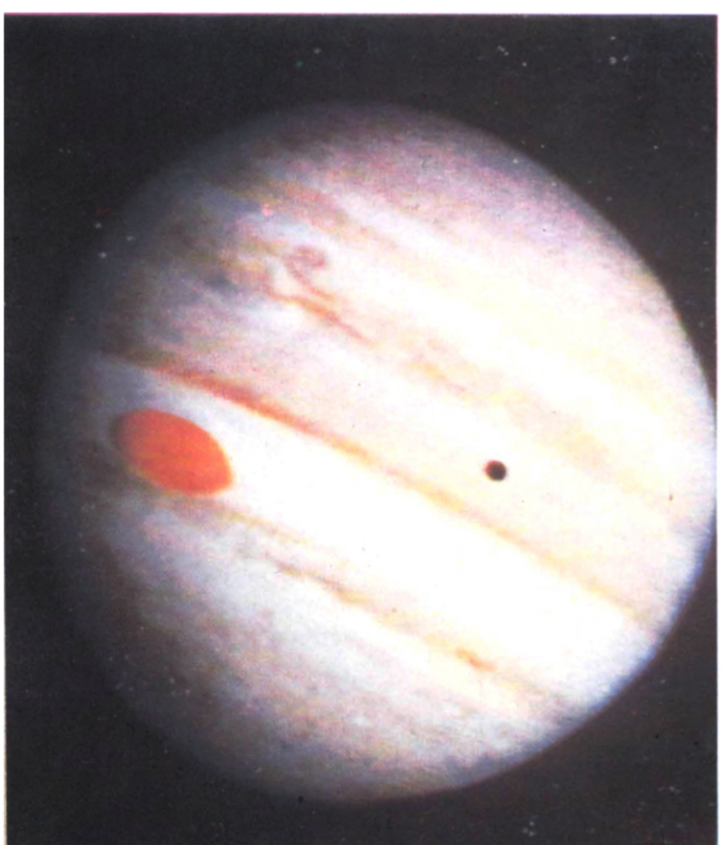
Is There Life on Earth?

The question, it seems, is not only superfluous, but also unusual. We do know that there is life around us here on Earth, but the question might be asked by some intelligent beings from another world in much the same way as we ask if there is life on Mars.

And how could they find an answer to this question?

We, of course, explore the physical conditions on a planet, compare them with terrestrial ones, and based on this we conclude whether there can be some sort of life there like on Earth. A scientific answer to the question can only rely on our experience of observation of life on Earth. Scientifically, it is only possible to perceive life as it is known here. We know, for example, that proteins – the basis of developed life – coagulate at elevated temperatures, but we can never know the limits of the adaptability of life or know if life on other bases is possible. We will now discuss the apparent traces of intelligent life. These, in principle, are spectacular from farther away than are intelligent beings themselves. Our best pictures of Mars, taken by the orbiters from less than 2,000 kilometres, in which we can distinguish features smaller than a few hundreds metres across, show no traces of intelligent activity. It is now undeniable that there is no intelligent life on Mars, but to what extent can the intelligent activity that we know occurs on Earth be seen from afar? This question interested the American scientist Sagan.

Jupiter from 2,500,000 kilometres with the shadow of Io (*Pioneer 10*).



He and his co-workers investigated in 1966 a wealth of pictures taken by meteorological satellites with a resolution up to 2-0.2 kilometre. It might be expected that the most apparent signs of intelligent activity would be the seasonal variations of the large, regular shapes of fields with agricultural crops. But the image contrasts were too small and any seasonal variation due to fields was masked by the seasonal variations of the angle of incidence of sunlight and aspect angle.

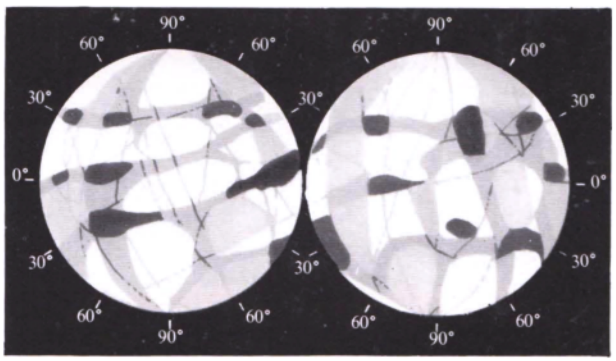
Large linear structures (roads, bridges, dams, wakes of ships on seas, or white condensation trails of aircraft) seemed more promising. These features were to be found in thousands of pictures without clouds.

Only one suggestion of a technological culture was found: a bright line of a section of highway. More doubtful were long lines, black and white, running parallel to each other above the cover of clouds. Obviously it was the condensation trail of a plane or its shadows. Another trace of human activity was the white network of clearings in forests covered with fresh snow.

The scientists concluded that if the hypothetical Martians had pictures of the Earth of the quality and quantity obtained by *Mariner 4*, the first probe that had passed 12,000 kilometres away from the Martian surface, they would not be able to find any signs of intelligent life on our planet.

The Giants Jupiter and Saturn

Beyond the orbit of Mars two giants among the planets – Jupiter and Saturn – are revolving ponderously. Jupiter is 13 times larger across than the Earth, while Saturn is "only" nine times larger. You can, if



Jupiter's moon Ganymede.

you wish, calculate yourself how many times the surfaces and volumes of Jupiter and Saturn are larger than the Earth's. Or you can look them up in any table of planetary data. A table would also give you the correct periods of revolutions of the planets about the Sun, but here it would suffice to say that Jupiter takes about 12 years to go round the Sun and Saturn takes around 30 years – so they are in no hurry. Not because they are huge (Jupiter is 300 times more massive than the Earth), but because their periods are conditioned by their distance from the Sun and the Sun's mass. This is due to the law of gravitation.

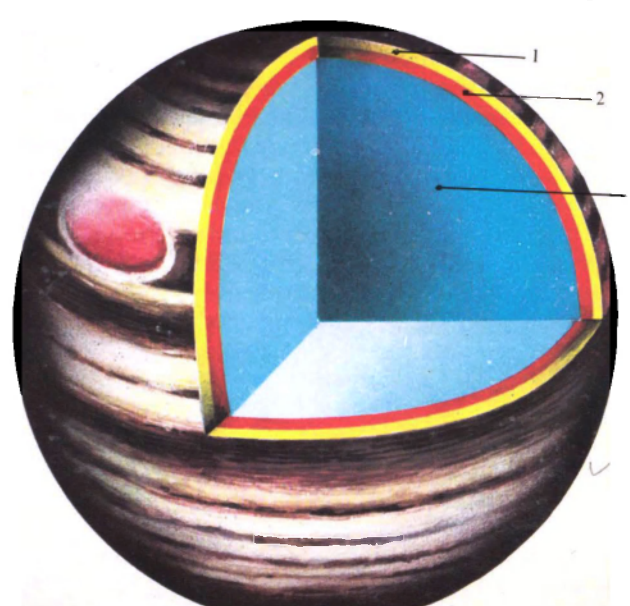
The prominent French popularizer of science Camille Flammarion (1842-1925) astonished his readers in the last century by the accuracy with which the distances or periods of the planets were known. But my own experience tells me that a reader is often not surprised by this accuracy of classical astronomy. When told, say, that the distance to the Sun is known to within 500,000 kilometres he might exclaim "You call that accuracy?" People do not always realize that always it is relative, rather than absolute, accuracy that matters. These 500,000 kilometres are only 0.3 per cent of the solar distance. You would hardly measure the length of your room with such an accuracy, although it is much easier than finding the Earth-to-Sun distance.

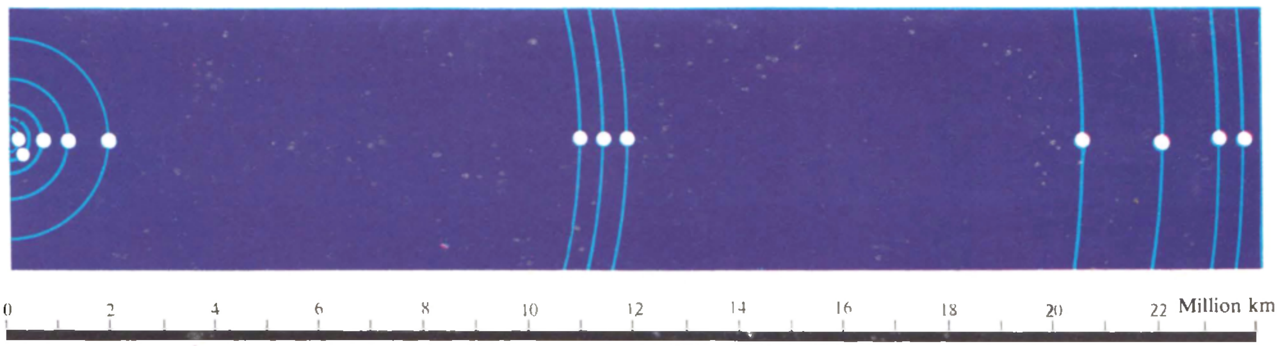
By the way, in recent years the Sun's distance, a unit astronomers use to measure distances in the solar system, has been found to within ± 0.001 per cent using radar techniques to determine the distances to the neighbouring planets.

I am not going to tire the reader with exact numbers (when they are known) which he would not remember anyway. On the other hand, there are reference tables where he can look them up. What is often amazing is not the accuracy of the findings of Earth-bound astronomy, but the frequently rough astrophysical evidence concerning some fact of nature whose existence was not even suspected or assumed. Later in the book examples of this will be more abundant.

The club of giant planets includes Uranus and Neptune as well, although they are far smaller than Saturn. The mean density of all the four planets is small (about that of water), and that of Saturn even lower (0.7 that of water). Studies of the polar compression of these planets owing to their rotation and the analysis of the influence of a planet's flattening on the motion of their satellites strongly suggest that their mass is concentrated at the centre much stronger than with Earth-type planets. The rather dense core of the Jupiter-type planets contains the bulk of the planet's mass. But their apparent volume is determined by the surface of the large opaque atmosphere compressed to the solid-like state at the bottom by the pressure of the upper strata. When we divide the known mass of the planet by this enormous apparent volume we arrive at the

A possible structure of Jupiter:
1 water-vapour and ammonia clouds; *2* gaseous hydrogen and helium; *3* solid hydrogen.





The distances of Jupiter's moons from the planet's centre. From left to right: Amalthea, Io, Europa, Ganymede, Callisto, Himalia, Lysithea, Elara, Ananke, Carme, Pasiphae, and Sinope.

mean density that amazes us by its smallness.

Some astrophysicists have lately estimated that the lightest gases, hydrogen and helium, account for up to 90 per cent of Jupiter's mass and that the temperature in the planet's centre may be as high as 100,000°C. At the same time, on the outside, due to heat losses, Jupiter may be as cold as we observe it from the Earth. Responsible for this picture is not so much the heavy chemical elements, as highly compressed hydrogen.

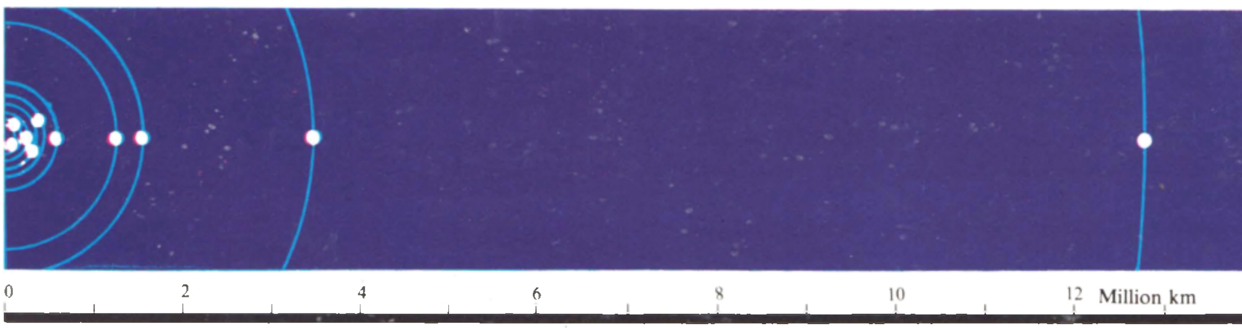
All the giant planets are surrounded with very dense cloudy atmospheres dominated by hydrogen and helium with some methane and ammonia, the latter freezing out the more the farther away from the Sun is the planet, i. e. the colder its climate.

The presence of ammonia and methane on the large planets is caused by low temperatures. These gases also form on Earth but they are quickly decomposed into components by sunlight, which is more intensive here. Moreover, for methane and ammonia to be formed in large quantities free hydrogen is required, which is almost absent in the terrestrial atmosphere. But the giant planets have retained their hydrogen, despite its volatility, since their beginnings. This was also helped by low temperatures and great pull of gravity at their surfaces.

All the four giant planets, especially Jupiter, rotate faster than the others. Furthermore their clouds rotate with different velocities at different latitudes, the fastest at the equator (for Jupiter the rotation period is 9 hours 50 minutes). Although the giant planets' density is small and they do not rotate like solid bodies, this does not imply that the planets themselves are hot liquid bodies as was sometimes assumed (at the time when physics rendered little assistance to astronomy).

When it became possible to measure planetary temperatures from their infrared emissions, it appeared that the temperatures of the giant planets were extremely low, and so astronomers had to drop the assumption that they are molten. Jupiter's temperature appeared to be about −140°C, and Saturn's about −155°C. Astronomers even succeeded in 1963 in measuring the temperature distribution over the Jovian disk. It turned out, unexpectedly, that its temperature is virtually the same throughout, the dark stripe of clouds being warmer than light ones. At the centre of the disk it is −141°C, being only several degrees lower at the morning and the evening edges and even near the poles. A phenomenon that has not yet been explained was observed in 1962. In the shadow of a satellite (and the eclipsed Sun) on the clouds the temperature appears to be 50°C higher than elsewhere. But within a quarter of an hour the temperature has returned to normal. Subsequent observations have not confirmed this, therefore further study is needed. A comparison of the measured thermal energy emitted by Jupiter with the computed energy received by the planet from the Sun shows that the former is larger. This leads to the conclusion that Jupiter has heat sources of its own.

A large reddish spot first observed about 80 years ago on Jupiter was once taken to be a lake of lava on its solid surface. The prominent Russian astrophysicist Fyodor Bredikhin (1831-1904) studied the Great Red Spot extensively as early as the 1870s. It was then supposed that the air currents emanating from it drive the clouds away from this site and make it visible. We now assume that it consists of some exceedingly light matter, solid rather than liquid, and is supported by the



The distances of Saturn's moons from the planet's centre. From left to right: Janus, Mimas, Enceladus, Tethys, Dione, Rhea, Titan, Hyperion, Japetus, and Phoebe.

sufficiently dense Jovian atmosphere high above its solid surface. The Spot measures 10,000 × 45,000 kilometres. Its solid nature is also borne out by the fact that it shifts as a single whole longitudinally. But still the nature of the Spot remains much of a mystery. In 1978-1979 the American probe *Voyager 1* flew by Jupiter and sent back close-ups of gigantic eddies near the Great Red Spot, traces of violent circulations in its atmosphere.

We have already mentioned that the rotation period of the Jovian clouds is several minutes a day less at medium latitudes compared with the equator.

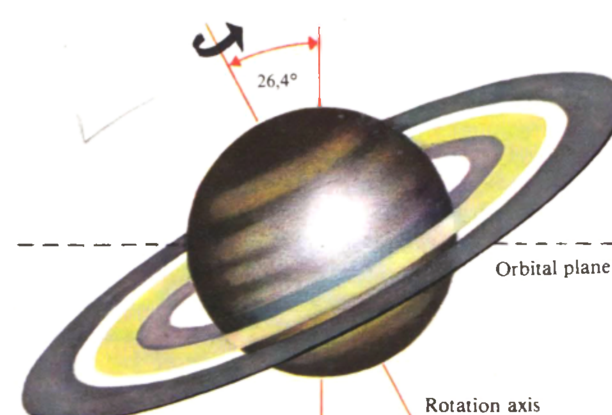
It has been found recently that sometimes the rotation speed, as determined from the inclination of ammonia absorption lines in the spectrum, is some two kilometres per second different from that derived from Fraunhofer lines which are due to the sunlight reflected by the Jovian clouds. Ammonia vapour is contained higher in the atmosphere of the planet. This suggests that winds blow there as it has recently been found in the terrestrial stratosphere, but with velocities 40 times faster than in our troposphere. It is not improbable that the different inclination of the various lines in the Jovian spectrum is caused by some other effect, rather than Doppler's.

At the edge of Jupiter's cloud cover the pressure is 1-2 atmospheres, but the density is about five times smaller than at the Earth's surface (like that at about 10 kilometres in our atmosphere). The clouds on Jupiter, stretched along its equator due to fast rotation, are easily seen in a small telescope, as is the marked oblateness of the planet, which is also caused by the fast rotation of Jupiter. The Jovian atmosphere may include water vapour and more complex molecules, specifically acetylene. The lines of Jupiter constantly change and the Great Red Spot alone, which in recent decades has lost its red colour, is the only permanent feature on the planet's disk.

Observations from *Pioneer 10* in 1973 and *Pioneer 11* in 1974 indicated that Jupiter has some sort of a hydrogen-helium corona. The upper cloud layer consists of ammonia spindrift clouds. From Earth-based observations, Jupiter radiates 2.5 more heat than it receives from the Sun.

Jupiter continuously produces radio emission of varying intensity and occasional "bursts". Radiation is seemingly produced by plasma waves in its ionosphere. It has been found that these radio bursts are associated with the orbital position of one of Jupiter's satellites, the one called Io. Io itself has a magnetic field and ionosphere and these interact with the Jovian

The axial inclination of Saturn.



ionosphere. Maximum emissions occur at Io's rises and sets for the planet's centre. Earlier still it has been found that the terrestrial artificial satellites produce noticeable ionization in the Earth's magnetosphere. It may well be that in Jupiter's system a similar thing occurs and a sharp increase in electron concentration is accompanied by radio emission. This strongly suggests that Jupiter has a powerful magnetic field and a radiation belt associated with the latter.

Further evidence of the Jovian magnetosphere came from the American *Pioneer 10* space probe, which in December 1973 passed within 130,000 kilometres of Jupiter's surface.

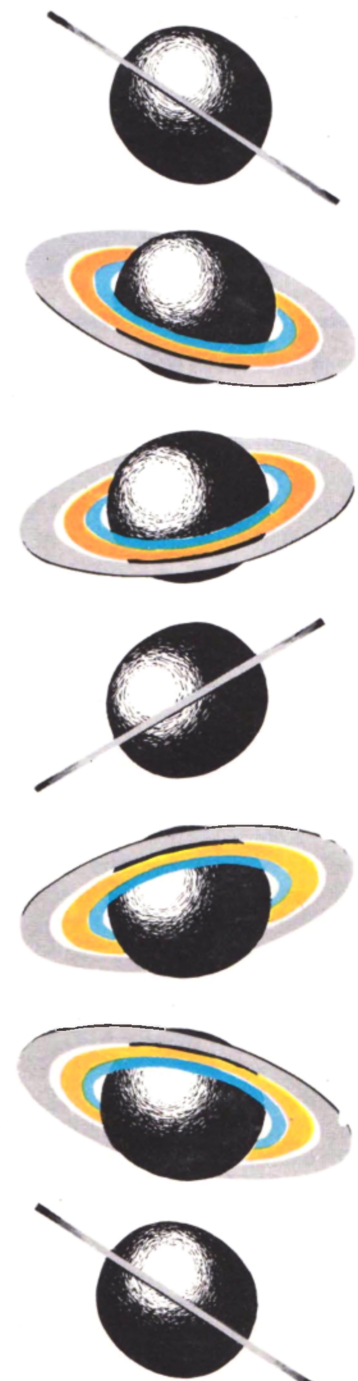
Pioneer 11 was launched later and flew by Jupiter in December 1974 whereupon it travelled towards Saturn's orbit. The Jovian magnetic field counteracts the interplanetary magnetic field produced by the solar wind. Where they meet the velocity of the solar wind halves, down to 200 kilometres per second, and its temperature increases from 10,000°C to 1,000,000°C. This bore out that Jupiter is surrounded by thick belts of energetic particles whose intensity is 10,000 times stronger than around the Earth. These belts are about 2,500,000 kilometres long. At the same time the instruments on *Pioneer 10* found, from the glow of the outer atmosphere's spectral lines, that the Jovian atmosphere was around 27 per cent helium.

The presence of a magnetic field on the Earth and the fast-rotating Jupiter and the absence of it on the slowly rotating Moon, Venus, and Mercury support the hypothesis that the fields are caused by rotation and currents in planets' liquid cores, if any.

Most of what has been said about Jupiter applies to Saturn and its atmosphere except that Saturn has been less studied. The straps in its disk are hardly noticeable and no radio emission has yet been found.

Jupiter has a retinue of 14 satellites, among which four large ones stand out distinctly. They were discovered by Galileo and you can even see them in the binoculars. All of them revolve around Jupiter facing it always with one side, just like the Moon revolves around the Earth and for the same reason. Their rotation has been retarded by tidal friction.

The larger of the four satellites, Ganymede and Callisto, are larger than Mercury, whilst Io and Europa are smaller than Mercury by a factor of around 1.5. Sometimes they pass between Jupiter and the Sun and then their shadows sweep across the planet's clouds,



The different positions of Saturn's rings.

and sometimes they disappear in Jupiter's shadow. These eclipses of the Jovian satellites played a very important part in earlier times. By observing them Römer of Denmark first established in 1675 the finiteness of the velocity of light and evaluated its value. In addition, comparisons of the times of eclipses predicted according to Greenwich time and their local observation times have long been used to determine the geographic longitudes of a locality.

The disks of the satellites are barely visible through even the strongest telescopes. Judging from their spectrum they have no atmospheres. The two satellites with the largest mass could have methane atmospheres but their proximity to the Sun results in their being heated sufficiently for any such atmosphere to have been scattered.

Although the same size as Callisto, Ganymede is almost three times brighter. This may be due to its being covered by a layer of white, frozen carbon dioxide with some other gases mixed in which there may be Ganymede's old carbon dioxide atmosphere, or rather the remnants of it is managed to retain.

The masses of the major satellites, as deduced from their mutual perturbations, are not known accurately. Provided the error is not too high, their densities are not large, but Callisto's relative density is estimated to be even 0.6. Obiously, it consists of frozen gases. Frozen gases seem to cover the surfaces of the other satellites as well, since they reflect sunlight far better than the Moon. *Pioneer 10* measurements indicated that the density of Io was 3.5 grammes per cubic centimetre (the same as the Moon's).

The remaining 10 satellites of Jupiter are 1/1,000 or even 1/100,000 as bright as the major satellites and are only visible on photographs taken using the strongest telescopes. They are rather small bodies and it is possible that they, like Mars's satellites, are former asteroids that "carelessly" came too close to mighty Jupiter and were caught by it. Evidence favouring this possibility is both their large distance to the planet and that four of them revolve in the opposite direction to the others. The eighth satellite moves in an unstable manner, and the Sun perturbs its orbit so much that it is difficult to observe without first having predicted its position for a given instant of time.

The fourteenth satellite is a new discovery. The 1978-1979 set of photographs sent back by the *Voyager 1* probe from a relatively close distance showed us the surface of Ganymede, Callisto, Io, and Amalthea. The first three are round and the pictures showed details not visible through terrestrial telescopes. Some very dark and some dazzlingly white spots stand out, obviously ice caps on their mountains, but the whole of their surface, it is expected, is covered with ice. At the time of writing radiomen and scientists are eagerly trying to catch the unexpectedly abundant information coming from *Voyager 1*. Unfortunately, a thunderstorm near the reception site interfered with the reception and several hours of important information continuously transmitted by *Voyager 1* were lost. Amelthea's configuration appeared to resemble half a cucumber. The surfaces of these moons were obviously subject to meteoric bombardment.

Saturn's retinue is fewer than Jupiter's, but the largest of its ten satellites is a little more massive than the major Jovian moons, and others too are not as small as Jupiter's other satellites. The most interesting Saturn's moons are Phoebe, which moves in a retrograde sense, and Titan, the only moon in the solar system with an atmosphere of its own – it consists of methane with, perhaps, some ammonia.

The atmospheres of Io and Ganymede are negligible in comparison with Titan's. All the bright Saturnian moons, with the exception of Titan, revolve around it facing it with the same side, as was indicated in 1971 by the measurements of their brightness.

At the Far Reaches of the Solar System

Orbiting at the outskirts of the solar system are Uranus and Neptune, belonging to the club of giant planets, although they are much smaller than Saturn, and also Pluto. Pluto is 40 times farther away from the Sun than the Earth, so that our luminary is only seen from Pluto as a dazzling star. In fact, it would be impossible to make out the Sun's disk from there, so small would the Sun appear.

There is not actually very much to say about Uranus and Neptune. Their atmospheres are vast, like Jupiter's and Saturn's, and their surfaces are fully masked by clouds. The clouds are arranged in obscure bands, just as on Jupiter. Because of low temperatures the ammonia in the hydrogen-helium atmospheres of Uranus and Neptune has partially frozen out. The temperature on Uranus was once believed to be $-180°C$, but observations of its radio emission at a

Uranus with its moons.

wavelength of 1.9 centimetres yielded an average temperature of $-100°C$ and $-170°C$ at 11 centimetres. This evidence seems to refer to different layers in Uranus's atmosphere. The temperature of Neptune appears to be 15-20°C lower. Since the first determination of the Venusian thermal radiation using radio techniques, the power of radio telescopes has increased more than 10,000 times. This has allowed us to measure radio emission from all the planets, save for Pluto. Radio studies of the latter await improvements in our hardware. Combined with infrared studies of the planets the findings of radio astronomy have provided important information. It was found that the temperature of the outer surface of the Uranian cloud blanket is only conditioned by the incident sunlight, and Neptune, like Jupiter and Saturn, radiates approximately 1.5 times more energy than it absorbs. Consequently, there are some very powerful internal sources of energy active there.

The rotation period of Uranus is 10 hours 50 minutes, whereas Neptune's period is known less accurately. Spectral evidence indicates that it is 16 hours with a possible error of an hour, though periodic variations in its brightness suggest it is 12 hours 43 minutes. The former figure seems to be more accurate. Interestingly, Neptune rotates in the direct sense, whereas Uranus in the retrograde sense, unlike all the other planets. Uranus's axis is inclined at 98° to its orbital plane, and so it rotates, as it were, lying on its side. This causes extremely sharp changes of seasons of the year, which on this cold and distant planet is 84 Earth-years long. But there, in the eternal cold, nobody is concerned about the seasons since that trinity of remote planets can in no way bear any life.

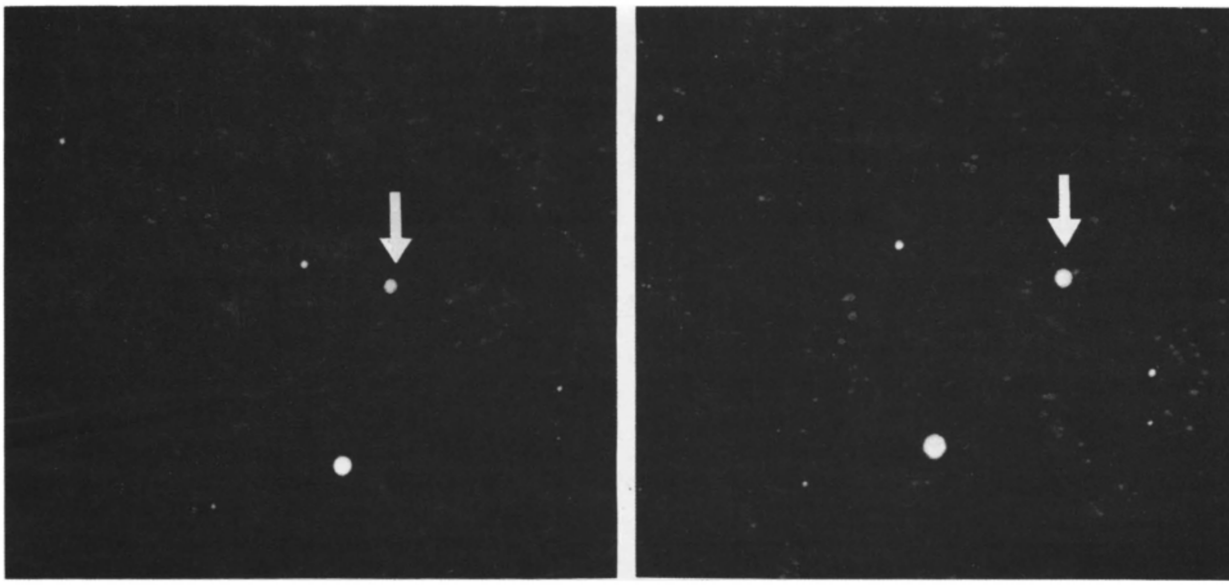
Pluto is a small planet, weakly lit by the Sun and so distant that only through the strongest telescopes it is seen as a light spot of the 14th or 15th magnitude.

By observing the regular changes in its brightness it was found that the rotation period of Pluto is 6 days 9 hours 17 minutes, but the period of brightness variation is not always an accurate indication of a planet's period of rotation and the two periods only coincide if a planet has a large and stationary spot (conspicuously light or dark). Clearly, we know virtually nothing about the physical nature of Pluto, but it is to be expected that it is similar to Mercury. However, with the terrific cold there, it might retain an atmosphere.

With the possible exception of Mars, all the distant earths are surprisingly unlike our planet. This diversity of planets and their atmospheres is another indication of the infinite variety of nature. It would be erroneous to think that everything in it is arranged after the same model. As a matter of fact, although the laws of nature are universal, the conditions under which they manifest themselves are different.

It may well be that beyond Pluto there are one or more other planets, but the effort that might go into a search for them among the wealth of faint stars is not justified and the time wasted would be better directed at more topical observations and measurements. Maybe you have read that Neptune was discovered "at the tip of a pen", in an arm-chair way. An interesting story is attached to discovery and the speculations about the planets beyond Pluto.

Uranus has five moons with orbital planes almost perpendicular to that of the planet, and they revolve around Uranus in the same direction in which Uranus rotates (retrograde rotation). In addition to the largest satellite in the solar system, Triton (diameter 4,000 kilometres and known since 1846), a second satellite, Nereid, was discovered in 1949. Nereid is at the limits of the capacity of the strongest telescopes and appears



The planet Pluto was discovered in 1930 by comparison of photographic pictures (different positions relative to fixed stars).

as an object of the 19.5 stellar magnitude. It is 17 times farther away from the planet than Triton and is only 300 kilometres in diameter. It is in direct motion, unlike Triton which is in retrograde motion.

Quite unexpectedly, an American astronomer named Christie, of the United States Naval Observatory, discovered in 1978 a satellite of Pluto as a slight elongation of the photographic image of the planet. The discovery was supported by pictures taken in 1965-1970, and later by pictures of three other observers. Its revolution period is 6.3867 Earth-days and its largest separation from the planet is $0''.9$. Its revolution period is equal to Pluto's rotation period, which may be the result of tidal friction between them. The discovery of the satellite enabled the mass and size of Pluto to be estimated. Its mass appeared to be around $^1/_{500}$ the Earth's mass, its diameter being about 2,600 kilometres, i. e. Pluto takes from Mercury the laurels of the smallest planet in the solar system.

The Ringed Planets

METEORITES IN SATURN'S RINGS. In the solar system meteorites are not only those that tumble down onto our planet, they also include those that perpetually circle round the distant Saturn.

The remarkable ring surrounding Saturn can be said to be a decoration not only of the planet but also of the entire solar system. For a man for the first time peeping through the telescope this, apart from the Moon, is, perhaps, the most fascinating view. We and Saturn owe this show to meteorites. It took more than two centuries to solve the mystery of this formation in our solar system.

Galileo, who turned his telescope at various celestial bodies and almost always discovered something unexpected that failed to comply with the Ptolemaic and medieval views of the Universe, was greatly perplexed by the "behaviour" of Saturn. In his crude telescope, which only gave blurred images at a 30-fold magnification, Galileo saw some "handles" on either side of Saturn, and he could not understand their nature. Actually, Galileo has seen the "ears" of the ring, i. e. the lateral parts of the ring and the dark gaps separating the ring from the Saturnian globe. These gaps led Galileo to assume that there is a smaller planet, something like a satellite on either side of Saturn. After all, he had already discovered four satellites accompanying Jupiter like a retinue. Why, argued Galileo, could not Saturn have something of the sort? It was impossible to make out the ring, which anyway would not have been thought of.

Which scientist, on the verge of a wonderous discovery, does not quiver, consumed by conflicting feelings: the jealous pride and fear of a blunder. To report or not to report, that is the question.

In Galileo's time scholars used the following way out. They promulgated the enciphered short anagram about the discovery, which could only be deciphered by the author himself. Once the author had checked his discovery definitively, he then unveiled his enigmatic report having thus retained the priority with him.

So Galileo also used this trick and published the following anagram:

Smaismrmielmepoetaleumibuvnenugttaviras

These letters, adequately rearranged, formed a phrase in Latin (the then language of scholars of all lands) that reported Galileo's discovery.

Anyone wanting to learn about Galileo's discovery can try to rearrange the letters till he happens to work out a sensible phrase. But then, there is no way of being sure that the same letters do not give an absolutely different phrase.

The mathematical theory of combinations enables the number of permutations (with repetitions) that can be made of these 39 letters to be worked out. Their number is

$$\frac{39!}{5!3!2!2!5!2!4!2!3!3!4!},$$

which yields a 35-digit number!

You probably wouldn't dare trying to decipher this anagram, but Kepler, the famous contemporary of Galileo's and one of the fathers of modern astronomy, did dare. Kepler's unbelievable patience, which also appeared of help in discovering his notable laws of planetary motion, was unparalleled. Kepler has thus deciphered Galileo's message (having discarded 2 letters) thus:

Salve, umbistineum geminatum Martia proles

Translated this is:

Salute, you twins, Mars's offspring

This was why Kepler believed Mars had two moons uncovered by Galileo.

Kepler thought that Mars, a planet that was situated between the Earth, which had one moon, and Jupiter, which had four moons (just discovered), should have just two moons. It is well known that the two Martian moons were indeed found... but only two and a half centuries later, whereas the family of Jovian moons is now known to be 14, nót just 4.

Alas, Kepler's labours turned out to be useless, for the anagram of Galileo, deciphered by him later, without two letters (included to make the anagram more difficult) said:

Altissimum planetam tergeminum observavi,

which meant

Highest planet, triple observed

Saturn, the most remote planet from the Sun among the then known, was called the "highest" planet. Galileo added allegorically that Saturn, so called in honour of the decrepit god, was supported on sides by two servants.

But soon the servants left their patron and several years later Galileo could no longer see them and so doubted his discovery. The trick, however, was that Saturn, at certain periods in its way round the Sun, turned so that its thin ring faced the Earth edge-on. It cannot then be seen through even the strongest telescopes, and several days before its "disappearance" it appears as a thin needle "piercing" the globe of Saturn. In medium-size telescopes, and even more so in the one possessed by Galileo, the ring becomes absolutely invisible, and seems to vanish.

This has repeatedly led to confusion. So in 1921 a number of provincial newspapers published reports by their zealous, but fairly ignorant, correspondents that "Saturn's ring had vanished", i. e. disintegrated, and some of them added of their own accord that "its fragments are flying to the Earth and a collision is imminent".

Saturn's ring vanishes every 15 years and in fact makes no trouble for us; on the contrary, it does us a favour by allowing us to see that it is extremely thin. We could obtain a scale model of Saturn's ring by cutting out a 30-centimetre ring from thin paper.

In 1966 the Earth crossed the plane of Saturn's ring three times during a nine-month period and the ring

twice showed its shadow side, remaining narrow. Another such occurrence was in 1980, and the next one will be in 1995. At times the ring opens up to show the whole of its surface, whilst at other times we see it "edge-on", its edges never appearing as a ring.

The true explanation of Saturn's ring and variations of its appearance was given, 50 years after Galileo, by the Dutch physicist Christian Huygens (1629-1695). And he, like Galileo, began with publishing the anagram

Aaaaaaa, ccccc, d, eeeee, g, h, iiiiiii, llll, mm, nnnnnnnnn, oooo, pp, q, rr, s, ttttt, uuuuu

Three years later, after a lot of checking and rechecking his initial assumptions, he reported the meaning of this enigmatic group of letters, which in Latin was

Annulo cingitur, tenui, plano, nusquam cohaerente, ad eclipticam inclinato

or translated

It is surrounded by a thin, flat ring not connected with the planet in any place and inclined at an angle to the ecliptic.

Later a dark gap was found in the ring breaking it up into two parts, the inner and outer, or Ring A and Ring B. This gap is known as the *Cassini division* after the astronomer who was the first to notice it. Then the narrower Encke's division and "crêpe ring", the innermost and weakly glowing, were uncovered. Therefore, they often speak about Saturn's *rings*, rather than one ring. It was found with time that Saturn's ring is not solid and not only in the sense that it is not a multitude of solid rings.

It was noticed several times that stars peep through the rings, almost unattenuated, thus suggesting a number of significant gaps through which a star shines as a lamp through a lattice window. In addition, one of Saturn's moons was once seen to sink into the rings' shadow. In the shadow of the outer bright ring the moon ceased to be seen, but in the shadow of the inner ring it only slightly decreased in brightness. This suggests that the inner ring is fairly tenuous, whereas the outer ring has few gaps.

Theoretical studies of the stability of the rings subject to the attraction of both Saturn and its more recently discovered satellites showed that a continuous solid or liquid ring would be destroyed by the attraction, it simply could not exist. Hence the rings had to consist of particles, small but numerous. This conclusion was reached by a number of scientists concerned with this difficult mathematical issue.

Definitive and unquestionable proof of the particulate structure of Saturn's rings is due to A. A. Byelopol'sky (USSR) and J. E. Keeler (USA). One of the founders of astrophysics, Byelopol'sky, was absolutely right in arguing that the final decision lies with spectral analysis. If the ring is continuous and rotates as a solid with the same angular velocity, then the linear velocity of the particles in the rings must vary in proportion to their distance from Saturn's centre. But if a ring consists of particles, then each of them must orbit Saturn independently of the others in its own orbit, being just like a small satellite of the planet. In that case its velocity must be determined by Kepler's laws, and the inner parts of the ring must revolve faster than the outer ones.

Doppler shift enables the velocity of a light source relative to an observer to be estimated. One edge of the Saturnian ring approaches us as it revolves, the other recedes with the same velocity. Byelopol'sky obtained spectra from different parts of the ring and found that the velocity of the inner particles in the ring (20 kilometres per second) is in fact far larger than the velocity of the outer particles (16 kilometres per second). Moreover, it appeared that the velocities of particles vary with their distance from Saturn's centre in exactly the same way as they should according to Kepler's laws, i. e. the squares of their revolution periods are proportional to the cubes of their distances from the planet's centre.

In 1934 Academician Shain of Simeiz Observatory, Crimea, USSR, again examined the Saturnian spectra but for another purpose. He was interested in the particle size within the Saturnian rings, which had escaped determination by direct methods. If the particles were chunks of material several kilometres across, at such enormous distances their disks would all the same be invisible. They appear just as glowing points that, being dense and numerous, merge for us into an apparently continuous ring.

A comparison of Saturn's spectrum with the Sun's

spectrum indicated no noticeable distinctions, so suggesting that the particles in the ring must be many times larger than the wavelength of light, i. e. much larger than one thousandth of a millimetre. If they were dust specks with sizes comparable with light wavelength, then they would scatter mostly the blue light in the spectrum of the incident sunlight. As a result, Saturn's rings would reflect blue light better than the rest of the spectrum, and they would be bluer than the Sun, and the blue end of the spectrum of the rings would be brighter than in the solar spectrum. Air molecules behave in much the same way as fine dust which is why the sky is blue. The blue part of the sky's spectrum is brighter than that of the solar spectrum, the air molecules scatter sunlight to make the day-time sky bright. The energy distribution in the spectrum of light reflected by Saturn's rings shows them to be common ice, rather than carbon dioxide ice. Obviously, the ring particles are covered by an ice layer or even consist of ice.

Unfortunately, this evidence gives us no indication as to the largest possible size of the particles within Saturn's rings: are they similar to the conventional meteorites that fall on Earth, or are some of them comparable to small asteroids? Variations in the brightness of the rings with the aspect presented to the Earth lead us to the conclusion that the majority of the particles cast rather long shadows, and hence are fairly large.

One of my students, M. S. Bobrov, was interested in the mysteries of the Saturnian rings before World War II. Later he compared the optical and other lines of observational evidence, and came to the conclusion that on average the particles are about one metre in diameter. His later work produced much smaller estimates – from 0.35 to 35 millimetres. But it was then found that the particles are slowly heated by sunlight, which is evidence in favour of the earlier conclusion of larger particles. Again, the larger particles are also favoured by the radar studies of Saturn, which were first carried out in 1973. No reflected signal came from the planet itself, whereas the rings gave stronger echoes than expected. This all led astronomers to conclude that the ring consisted of angular chunks a metre or more across.

The internal crêpe ring, according to Shain, appears to be bluer, i. e. it must partly consist of fine dust particles comparable in size with the wavelength of light.

Saturn's ring is so wide that the Earth, whose diameter (12,756 kilometres) is five times smaller than the ring width, could comfortably slide along it. Of the three constituent rings, the middle one is the brightest and densest (Ring B), its width being 26,000 kilometres. The Cassini division, which separates from it the outermost Ring A, is 5,000 kilometres wide, and Ring A is 16,000 kilometres wide. The semi-transparent crêpe ring C glows weakly and is transparent enough for the planet's surface to be seen through it; the width of this ring is 18,000 kilometres. Now imagine, these enormously wide rings are only 1.5-3 kilometres thick!

The total mass of the ring can be estimated from the theory that explains its stability. The ring's mass is no larger than a quarter of the Moon's mass, and is most likely far smaller, as it has been found from observations of perturbations suffered by Saturn's satellites from its pull. The rings and satellites produce mutual perturbations. See how much more sunlight is reflected by this small mass, than by the same mass collected into a ball! If a quarter of the lunar mass were turned into meteorites and arranged in a circle round the Earth, it would illumine us thousands of times stronger than now.

It would be fascinating to enjoy the sight of the rings from Saturn itself. Alas, we would be in for a disappointment because from the poles to 64° latitude on Saturn the rings are not seen at all, since they are blotted out by the curvature of the globe, and only in the equator regions of the planet, between +35° and −35°, we could see the wide expanse of the rings. Even here, however, they would be seen at an angle of 12° or less, rising above the horizon like a rainbow, and from the equator itself they could only be seen edge-on, as a bright, but extremely narrow stripe passing through the zenith and dividing the skies in two. If we also take into account that just one side of the ring is illuminated and the other will be dark, we must conclude that in either of Saturn's hemispheres the rings could only be seen during half a year (of course, a Saturnian half-year which equals 15 Earth-years). The rings in a given hemisphere of Saturn would mostly be seen by day, which detracts from the beauty of the show, and by night part of the ring would be covered by the shadow of the planet itself. Finally, if we remember that Saturn is always covered by clouds which completely envelope its atmosphere, we will be led to conclude

that we would not see any rings from Saturn's surface. So if you want to get a better view of the particulate rings of Saturn you had better not stay on Saturn. The best way to see them is from one of the planet's satellites. Again they could only be seen from a small angle, almost edge-on.

It may well be that once one of the Saturnian moons, closest to it, that dared to approach yet closer was punished for its audacity by being fragmented and turned into the particulate ring.

The French mathematician E. Roche (1820-1883) proved in 1850 that a liquid satellite of a planet that approached within some limiting distance must be torn into pieces by tidal forces. For a satellite with a density equal to that of the planet the "Roche limit" is 2.44 planet's radii. The closest of Saturn's satellites, Mimas, is separated from the planet's centre by 3.11 radii of the planet, and the outermost edge of the ring by 2.30 radii. Thus, the Saturnian rings lie wholly within the Roche limit, inside the zone forbidden for the satellites that want to continue as single wholes. This supports the validity of Roche's theory and the possibility that the rings were formed at the expense of a disintegrated satellite that either had begun to form in this region at the origin of the solar system but failed to develop into a satellite, or it had been perturbated into this zone by external forces. It should be noted, however, that for a continuous solid satellite the Roche limit is far smaller than for a liquid one.

The year 1966 saw the discovery of the widest outer ring, named Ring D. It is twice the size of the ring system known until then, but the ring is extremely tenuous. The satellite Janus discovered in 1966 and the more distant satellite Enceladus move inside Ring D as in a medium with resistance, whereas Mimas only intersects Ring D, because its orbit is inclined to the ring's plane. This latest evidence agrees with our last statement concerning the Roche limit.

The gaps in the rings have been formed due to disturbances of the particles by the attraction of Saturn's satellites. Where the revolution period of the particles is compatible with the period of one of the inner satellites, the disturbances are especially large and make the orbits of the particles unstable. In places where the revolution period of particles is a $^1/_2$, a $^1/_3$, or a $^1/_4$ of that of the nearest satellites, the particles do not stay there long with the result that these regions are almost free of their orbits, and hence of the particles, and so we see there gaps. Similar gaps are observed in the orbits of asteroids, in those places where their revolution periods would be compatible with Jupiter's period.

The Cassini division in the rings corresponds to revolution periods equal to $^1/_2$ of Mimas's period, $^1/_3$ and $^1/_4$ of the revolution period of the two next satellites. Analogous ratios are inherent in other gaps in the rings.

RINGS OF OTHER PLANETS. Professor Vsekhsvyatsky (of Kiev) published in 1962 his work arguing that Jupiter, too, should have particulate rings, if only extremely tenuous ones. But the article did not attract any attention of the astronomic community.

On 10 March 1977 two teams of workers observed electrophotometrically the predicted occultation of a weak star by Uranus. It was expected that just before it goes out completely being blotted out by Uranus the star would a bit earlier decrease in brightness due to the absorption of its light in Uranus's atmosphere and would thus give data on the properties of the atmosphere. Much to their confusion, however, the star "winked" 35 minutes earlier, and then again, and again. The star winked again after the planet had passed it. The phenomenon was observed both by a ground-based observatory and by an orbiter equipped by an identical electrophotometer. It was inferred that the light of the star had been successively attenuated by narrow layers of matter forming rings around Uranus, the rings being so tenuous that they cannot be seen directly. Just as with Saturn, they lie within the Roche limit where no large solid or liquid satellite could exist.

Astronomers then paid attention to Jupiter and found near it some extremely tenuous rings unseen by the eye. So, Vsekhsvyatsky's prediction was borne out in such an unexpected manner. According to Vsekhsvyatsky, the rings are due to powerful volcanic ejections from the surfaces of the satellites of large planets.

Are There Any Other Planetary Systems?

Are there planetary systems near other suns and where are they? It is impossible to see through the telescope a planet the size of the Earth even about the nearest star (α Centauri). If the hypothetical planet lies at the same linear distance from α Centauri as the Earth

from the Sun, then its orbital motion will, if only at times, bring it away from the star by 0″.75. A large telescope can resolve stars of similar brightness on clear nights. However... not only the Earth, but even such a giant planet as Jupiter will (theoretically) appear in reflected light as a star so weak that it will not be seen even through the largest telescope. Our Earth is not therefore very noticeable from a distance.

In the sequel to this volume we will see that single stars travel linearly through space, whereas their projection on the celestial sphere appears as a large arc. Also, we will learn that stars often form pairs. The members of a pair revolve around their common centre of gravity, which moves in space along a straight line, each of the sister-stars describing on the celestial sphere a wavy line. Sometimes one of the stars in the pair is not large and the apparent path of the stars gives us a clue to their respective sizes. The component stars, as a rule, have comparable masses. Some of the near stars display wave motions. This strongly suggests that each of them has an unseen satellite, whose mass is only a few times larger than that of our Jupiter.

The Soviet astronomer Professor B. V. Kukarkin indicated that in actuality the observed period of oscillations of such stars may be caused not by one massive planet, but by several smaller ones having different revolution periods, i. e. by an entire planetary system. Their combined action may produce the same effect as the attractive pull of one massive planet.

It is quite probable that, although we do not see them, we have already found several planetary systems.

We do not yet know definitively the way in which planetary systems form. Maybe not all stars acquire planets. It would, however, be highly improbable, with the enormous multitude of Sun-like stars in existence, for none of them to have planets. It would be more reasonable to suppose that most of them are endowed with a planetary family.

Some enquiring minds have always maintained that it would be highly unlikely for life in the Universe to be solely confined to our Earth. Giordano Bruno was sentenced to the stake for his teaching that there were multitudes of inhabited worlds and his other "herecies". The Holy Synod of the Russian Orthodox Church banned the second edition of *Talks about the Multitude of Inhabited Worlds* by Mikhail Lomonosov.

Dialectical materialism teaches that the infinite Universe harbours an infinite number of inhabited worlds and that life will unfailingly spring up where favourable opportunities present themselves.

The example of the solar system tells us (and this is also confirmed by theory) that not all planets are suitable for life, especially for intelligent life. We will be looking at this issue in the section "Is Any Communication Possible with Extraterrestrial Civilizations?" in the sequel volume "The Universe Beyond the Solar System".

2. The Little Planets

The Planet Hunters

"Nature abhors a vacuum", asserted many scholars in antiquity, although they found it difficult properly to explain where they got this from. By this assertion they "explained" why it is that the liquid in a barometer rises up a tall glass tube from which the air had been pumped out. Very likely a similar idea about the abhorrence of a vacuum caused scientists at the end of the 18th century to search for a planet in the "gap" that yawned between the orbits of Mars and Jupiter.

In 1766 in Bonn Johann Titius (1729-1796) published a book called *The Contemplation of Nature*, in which he called attention to a regular increase in the distances of the planets from the Sun and to a gap between Mars and Jupiter. This idea was picked up by Johann Bode (1747-1826), who declared that the vacant place in the solar system was occupied by a planet that he, one could say, simply invented. It should be noted that Neptune and Pluto which were discovered later do not fit the pattern noticed by Titius since in reality the law is more complex. The law made a great impression on people at the time, and in our time it may be explained theoretically. Titius's and Bode's assumptions had nothing in common with scientific prediction based on the law of nature, they were more closely related to the mystical ideas of the Pythagoreans about "sacred" numbers and figures.

The astronomer Franz Zach (1754-1832) believed so much in the existence of a planet between Mars and Jupiter that for fifteen years he attempted to organize a search for it, although as we have seen the assumption that a planet existed there was based on an obscure relationship among certain numbers. In September 1800, Zach succeeded in arranging the search with five astronomers who had gathered at a conference. They formed, as Zach jokingly observed, "a detachment of celestial police" with the aim of "tracing and apprehending a fugitive subject of the Sun." For this search, the sky along the constellations of the zodiac, through which all the planets move, was divided into sectors and each observer was allotted a sector of the sky. It was decided to offer the one unallotted sector to the Italian astronomer Giuseppe Piazzi (1746-1826). No sooner had they resolved to inform him of this, and had started their laborious, and what promised to be a long, if not hopeless, job, that they received news that the fugitive had been caught. It had not happened as a result of a long search, but quite by accident.

On the first night of the 19th century (1 January 1801) Piazzi in Palermo was diligently making his systematic measurements of stellar coordinates for the catalogue of star positions he was compiling. The next night, as he was repeating his observations as a check, Piazzi noticed that one of the faint stars he had observed (seventh magnitude) did not have the same coordinates he had noted the night before. On the third night he discovered that he had not made a mistake, the little star was moving slowly. Piazzi decided that he had discovered a new comet. True, comets, as Piazzi knew, were "fuzzy luminaries" which is what their name means in Greek; a comet is a body with a nebulous appearance, sometimes having a nebulous tail. "Perhaps it is an unusual comet", Piazzi decided, "such that hasn't been seen before."

Piazzi followed it carefully for six weeks until sickness confined him to his bed and interrupted his observations, from which Piazzi himself was unable to derive the orbit of the body he had descovered in space.

His illness over, Piazzi again began to spend his nights at his telescope, but now he could not find his comet again. Its continuous motion had carried it far from the place where he had last seen it, and it had become lost among other equally faint stars that at that time had not yet been entered into the star maps.

Thus, without completing his discovery, Piazzi had to send off letters to other astronomers with a description of his observations and a request to search for the body he had found and lost.

By the time the post had delivered these letters to other countries, it became quite hopeless to search for the new object. It couldn't be found, and the observations that had already been made were not sufficient to permit the orbit of the body to be computed by existing methods or to predict its apparent position in the sky in the future. It is difficult to guess how many years a new advance in astronomy would thus have been delayed, but here one discovery came to the aid of another. Theory aided practice. The observers were assisted by the mathematician Karl Friedrich Gauss (1777-1855).

Gauss was then only 25 years old, and although he had many plans, hopes and interests, he still had not

decided to which field he would ultimately devote himself. So he devoted his spare time to higher mathematics and astronomy.

Even before the events described here Gauss had found a general method for computing the orbits of bodies from just three observed positions in the sky. Gauss's new method had not had any application yet, and the new discovery presented a first and excellent opportunity.

Gauss immediately set to work calculating, and in November he published the elements of the planet's orbit and its future positions in the sky–where the planet ought to be seen from the Earth.

September 1801 had arrived when Piazzi's object, until then hiding in the Sun's rays, was supposed to emerge again and become accessible for observation–if they could manage to find it. Alas! The patience of the observers, who were eager to exploit Gauss's help as soon as possible, was to be tried severely. Rain, snow, fog and clouds seem to have conspired to prevent the search for the lost body, and it was only on the last night of that same year that the sky cleared at last. Untroubled by the approaching festivities to greet the new year and the freezing weather, Zach threw himself into the search "hot on the trail", and on the following night, on the anniversary of Piazzi's discovery, the fugitive was found. Its shift among the stars over a two-day period "nailed it red-handed". Moreover, on that same night it was also discovered by Olbers.

Gauss's calculations showed that Piazzi had discovered not a comet, but a planet revolving around the Sun between Mars and Jupiter. Who but Piazzi should have the first word in naming the newly discovered member of the planet family? And Piazzi wanted to call it Ceres, the patron goddess of Sicily during Roman times. Thus Piazzi paid tribute to the locality where he had successfully carried out his scientific work, and at the same time he "upheld tradition" since he took the name of the planet from the same pantheon of the gods of Roman mythology from which the names of the other planets were all drawn in antiquity.

The story of the naming of Ceres is one example of a possible answer to a question that naive people sometimes ask: "We grant that it is possible to measure and work out the size, distance and temperature of celestial bodies, but how pray tell, have you found out their names?" In the same way as parents find out the names of their children!

The discovery of the eighth planet was followed by a number of other discoveries, and in our day, as we shall see, we almost regret that there have been so many of these discoveries.

A Chain of Discoveries

Ceres was the object of constant attention, and while observing it astronomers made a thorough study of the location of the faint stars near its path. On 28 March 1802, Olbers noticed a new star not far from where Ceres had just recently been visible, and after only two hours he was sure that it had moved with respect to its neighbours. It looked like the discovery of still another planet, and again Gauss showed that this was indeed the case. It was especially surprising that the orbit of this second faint planet proved to be extremely close to that of Ceres. Instead of there being one planet lacking between Mars and Jupiter it turned out that there were two: "We were broke, and all of a sudden we had a tuppence." The second planet was called Pallas (an epithet of Athens, the Greek goddess of war, victory, wisdom, and science).

In former times there were few observatories and only a few people engaged exclusively in astronomical research, their work was poorly paid. About half the outstanding scientists of the 17th and 18th centuries studied science in their spare time, time taken from their other occupations that supported them.

For example, the famous astronomer Bessel began his career as a clerk, and Lassel, the discoverer of Neptune's satellite, was a brewer. Of the comet investigators, Swift was a tinsmith and Tempel a lithographer. Schreter, a planet researcher, was a court functionary; Herschel started his career as a musician; Schwabe, the discoverer of the periodicity of sunspots, was a pharmacist; Hall, who discovered the satellites of Mars, was a carpenter; and the investigator of the little planets, Olbers was a practicing physician.

By snatching time from sleep Olbers observed comets and became an authority on their orbits. In 1779, while he was attending a sick friend who like Olbers himself was a medical student, he hit upon an important simplification in the calculation of these orbits. Scientists sometimes get happy thoughts unex-

pectedly, even in the most inappropriate situations – in a streetcar, during a concert intermission, or even in a store. Immersed in his studies the scientist constantly tries to snatch every free minute for thought, and of course Olbers had this happy idea while on duty not by accident, but as the result of a long period of thought during the preceding months. To the question, "How did it occur to you?", the truest and shortest answer would in most cases be "I thought about it constantly." The new method simplified the work of those who calculated cometary orbits and speeded up the calculations.

The combination of rigorous thought and a certain amount of imagination is useful, and imagination nudges an investigator toward new discoveries. So Olbers expressed the bold idea that the place in the solar system that some people had assigned to just one planet and where now there were two little ones, had at one time really been occupied by a single planet. The two planets that had been discovered there, according to Olbers, were fragments of it that had been formed sometime in the past by some sort of catastrophe. He believed it very likely therefore that there were not just two, but many of these fragments, and there was good reason to hunt for the others.

If a planet located in between Mars and Jupiter had at some time exploded into pieces then the orbits of all the resultant fragments must pass through the point in space where the explosion had taken place. This is a well known law of mechanics, which must hold true even in space. If this were true, why grope around over a large area of the sky looking for new planets when it would be easier to intercept them when they passed through the points where the orbits of Ceres and Pallas intersect. This was the practical conclusion from Olbers's supposition and it could be called the "working hypothesis".

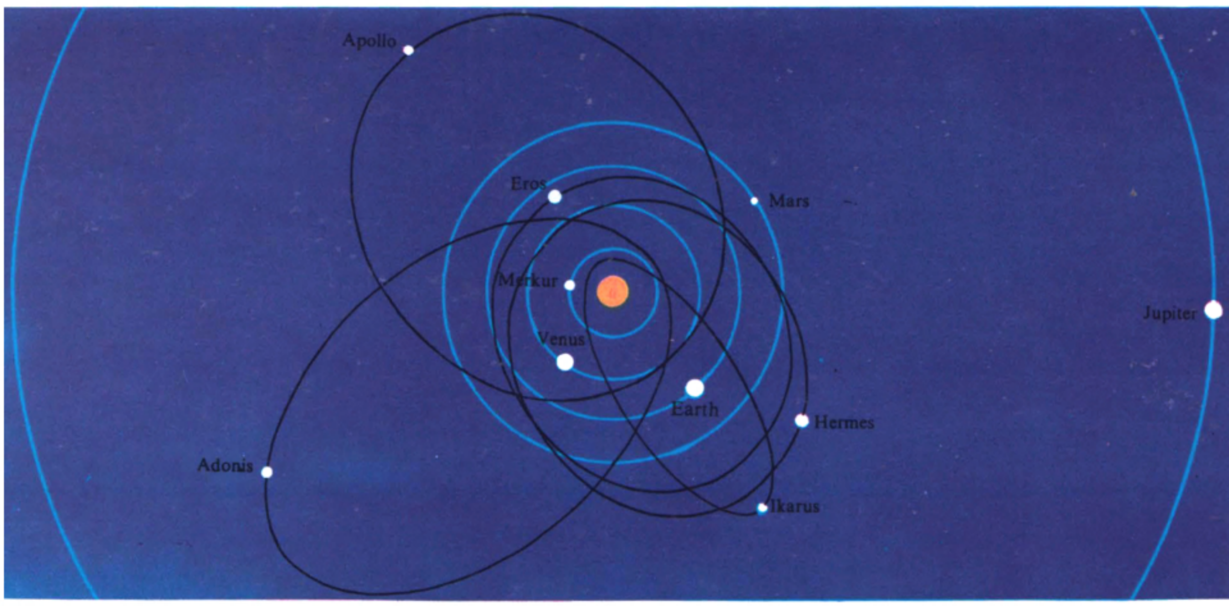
A "working hypothesis" is an assumption we put forward tentatively to account for a newly discovered fact of nature, although the fact itself is insufficiently studied for the assumption to be on firm ground. The working hypothesis, while not making any pretense at rigour, does for a while give some explanation for the evidence at hand and indicate paths of research to the investigators. Further research then progresses not blind but in a definite direction, and with the main purpose of verifying the validity of the hypothesis. After all, a hypothesis suggests some corollaries, for example, that a particular phenomenon should occur. All attention then turns to the question as to whether or not these phenomena are in fact there. If the hypothesis does not hold, a new one is advanced in its place, one which is now much improved because the work on verifying the first hypothesis has led us to a deeper understanding of the facts discovered and has added a fresh piece of evidence.

For three years Olbers himself patiently watched for new planets in the constellation Virgo, the point where the orbits of Ceres and Pallas as seen from the Earth intersect. His labours were rewarded in 1807 by the discovery of Vesta. But earlier, in 1804, Harding had discovered a little planet named Juno in the constellation Cetus, which was where the second point of intersection of the orbits had been located.

It seemed thus that the hypothesis was confirmed, and that the orbits of the four fragments intersected almost at the same points. When you think about it, however, Olbers's hypothesis would only be true if the catastrophe to the vanished large planet were fairly recent. In fact, if the event had occurred long ago, the gravitational attraction of the large planets would have altered the orbits of the fragments so greatly and in such different ways that they could in no way be expected to continue to intersect each other at the same points. The planets discovered since then (all in the same place, between Mars and Jupiter) by no means pass through the points where the orbits of the first four intersect. The initial impression that Olbers's assumption was valid proved to be based on a coincidence. All this became clear, however, only quite a long while after Olbers had found the fourth planet.

By the time all those who had taken part in the discovery of these planets had passed away, observers had still not seen the fifth planet. It was only in 1845, almost 40 years later, that it was found. It was discovered by a retired postal official, Encke, whose patience was truly astounding. For 15 long years, night after night, he searched for the fellow-travellers of Ceres, and each new night brought disenchantment but failed to weaken his enthusiasm. Two years after his first success he discovered another planet and soon discoveries of these planets began to be made all the time.

All the planets that have been found between the



The orbits of some asteroids.

orbits of Mars and Jupiter* have come to be known collectively as the minor planets, or asteroids, which is the Greek for "star-like". Indeed, even in the most powerful telescopes these planets appear as stars, so small they are. Smallness in astronomy is, of course, a relative concept, but in comparison with the other planets the asteroids are small indeed. Ceres, the largest, is only 1000 kilometres across, and in volume it is as much smaller than the Moon as the Moon is smaller than the Earth. Pallas has a diameter of 600 kilometres, Juno is 250 kilometres across, and Vesta 540**. Only these four can be seen as disks and only by using the largest refractors in the world. The diameters of these planets can be measured, but no details can be seen on them. The diameters of the other asteroids are far smaller and are estimated from their brightnesses. With a given surface reflectivity and a given distance from the Sun and from the Earth, the apparent brightness of a planet is proportional to the square of its diameter. Assuming that the surface of asteroids reflects about 7 per cent of the incident light, as do other bodies like the Moon, it is possible to estimate the approximate size of these tiny planets. The smallest asteroids now known have diameters of about 1 kilometre and could easily fit into one of London's parks. Variations in their brightness suggest that they are not spherical, but resemble irregular fragments, different sides of which reflect light somewhat differently. The observed rapid fluctuations of their brightness seem to be caused by their rotation about their respective axes (whereby they turn their brightest sides towards us at some times and at other times their fainter sides, which differ in size a bit).

Since the diameters of the four largest asteroids have been measured directly, it is possible to determine their reflectivity. For three of them it is between 5 and 16 per cent, i. e. close to the reflectivity of the surfaces of the Moon, Mercury, and terrestrial rocks. But Vesta reflects 23 per cent of the sunlight it receives, which is what is generally found with bodies that can be called white.

Reflectivity, size, and distance from the Sun (which generally don't change markedly) and the distance from the Earth (which does change significantly) determine the apparent brightness of asteroids. At

---

* In 1977 an asteroid was found with the orbit that almost completely lied between the orbits of Saturn and Uranus. It was named Chiron after a hero of Greek mythology, the centaur, part horse and part man. Its size is estimated to be 150-650 km.

** So far the sizes of more than thirty of the small planets are known, Juno being the 15th largest.

opposition, when they are closest to the Earth, Vesta is the brightest and just at the limits of naked eye visibility. The other bright asteroids are only visible in strong binoculars as stars of the seventh magnitude or fainter. Most asteroids are only visible in powerful telescopes and on pictures taken by large astrographic cameras, appearing just like starlets, without disks.

The Farther into the Forest, the More Firewood There Is (*Russian proverb*)

The smaller asteroids are and the fainter they are, the more of them there turn out to be, and therefore as time goes on fainter and fainter asteroids are discovered. For example, most asteroids discovered in 1930 fell into the 14th magnitude class, while the asteroids of 1938 were approaching the 15th magnitude.

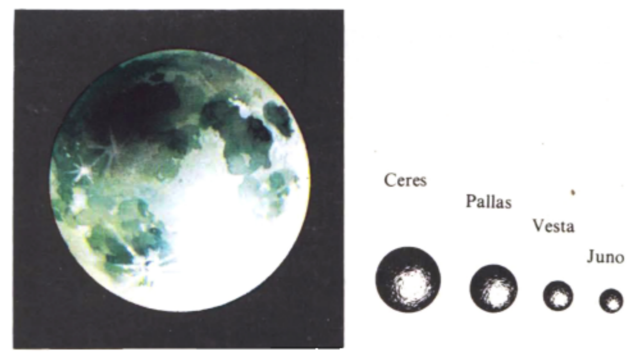
Photography is now the only method used to catch these minor planets, and even at the close of the last century when it was first used it immediately demonstrated its advantages over the visual telescopic searches of the 19th century.

To tell a faint asteroid from a star it is necessary to be sure that it is moving among the stars from night to night. If several nights are cloudy the suspected little planet may be lost.

On a photograph, when the camera tracks stars, they come out as little points, while a planet even during a one-hour's exposure shifts so much that it comes out as a short line, thereby giving itself away immediately.

If the asteroid is faint its trail may not be recorded on the film, and in order to catch the faintest asteroids the following method was invented. The clockwork is set so that the camera moves about in the direction the asteroid is expected to follow among the stars, and at the expected speed. With this method the images of the stars are blurred, coming out as short little lines, while the light from a faint asteroid falls all the time on almost the same spot on the plate and therefore has a noticeable effect on it. The photograph of an asteroid comes out to be almost a point.

After an asteroid has been found on a photograph it is necessary to make certain that it is a new one, and for this it is necessary to determine accurately its apparent position among the stars, its coordinates on the celestial sphere, and then to compare them with the ephemeris, i. e. with the apparent positions, calculated in advance, of asteroids whose orbits are already known.



The largest asteroids and Moon shown to scale.

Now, after making certain that you have discovered a new member of the solar system, you must obtain at least two more determinations of its position in the sky, and if possible, not on consecutive days, or else it will be impossible to calculate its orbit accurately enough. How many times has cloudly weather prevented an observer from following an asteroid and caused him to lose it because he had not managed to obtain the required observations of its position. A discovery that had seemed within his grasp slipped right between his fingers the way Ceres almost did once.

Until the orbit of a planet has been determined with sufficient accuracy from a large enough number of observations, the planet is assigned neither number nor name, and it is not considered worthy of becoming a member of the solar system.

Many such planets have been found and lost, then found again by somebody else and lost again, and therefore not every newly discovered planet receives a name immediately.

A great many minor planets (over a hundred) have been discovered at the Simeiz Observatory by G. N. Neuymin, S. I. Byelyavsky and others. When hunting for minor planets at Simeiz, two identical cameras are used to photograph the same region of the sky simultaneously, thereby eliminating the chance of any mistake.

After computing the orbit and having the new planet recognized, the observer has the ligitimate pleasure of giving it any name at all. However large the reserve of gods in the Greco-Roman pantheon was, it was not sufficient for naming the asteroids. The style

of mythological names could not be maintained, and people began to name asteroids as best as they could, just so long as the name had a feminine ending. Asteroids were named in honour of wives, daughters, and perhaps, even mothers-in-law.

Masculine names were only given to the most unusual asteroids, as an exception, in order to separate them from the multitudes of others.

Several minor planets have been given names in honour of the cities where they were discovered, like Moskva (No. 787) and Simeiza (No. 748). There is the asteroid Vladilena (No. 852) in honour of Vladimir Ilyich Lenin, and Morozoviya (No. 1210) in honour of Nikolai Morozov (Russian scientist and revolutionary), and others. Amundsenia, Piazzia, Olbersia, etc., were named in honour of scientists.

The following table gives a chronological picture of the discovery of asteroids. In the 1970s the numbered asteroids totalled 3,000. The table clearly shows the effect of photography, which came on the scene at the end of the 19th century and raised the annual number of discoveries.

The Discovery of Asteroids

Years	Discovered	Numbered	Total numbered
1800-1809	4	4	4
1810-1819	0	0	4
1820–1829	0	0	4
1830-1839	0	0	4
1840-1849	6	6	10
1850-1859	47	47	57
1860-1869	53	52	109
1870-1879	105	102	211
1880-1889	80	76	287
1890-1899	264	165	452
1900-1909	776	213	665
1910-1919	788	249	914
1920-1929	1262	202	1116
1930-1939	2799	373	1489
by 1962	–	–	1650

The size and mass of the asteroids is in some measure proportional to their brightness (reduced to equal distance from the Earth and the Sun), and therefore the distribution of the asteroids according to what is called their absolute brightness (i. e. the brightness an asteroid would have at a distance of one astronomical unit from the Earth and the Sun) describes their distribution according to mass (if their reflectivity is taken to be the same).

It is possible to estimate the proportion of asteroids at a given distance from the Sun that have not yet been discovered by studying the discovery statistics. In general we can conclude that there are a total of 530 asteroids brighter than the 9th absolute stellar magnitude. The number of fainter asteroids increases about 2.7 times as their brightness decreases by a factor of 2.5. With the largest telescope in the world, if it were possible to use it solely for finding minor planets, it would be possible to find 30-40 thousand of them. There may be hundreds of thousands of even smaller and fainter asteroids and according to estimates by S. V. Orlov, there may be around 250 million asteroids with a diameter under one kilometre, but in fact they add very little to the total mass of all the asteroids, which according to all the data does not even exceed the mass of the Moon.

These appallingly large number of little planets move in all the possible orbits between Mars and Jupiter, and their paths are so intertwined that if we were to make a wire model of their orbits in the form of rings, there would not be a single ring that could be pulled out of the model without pulling all the others after it.

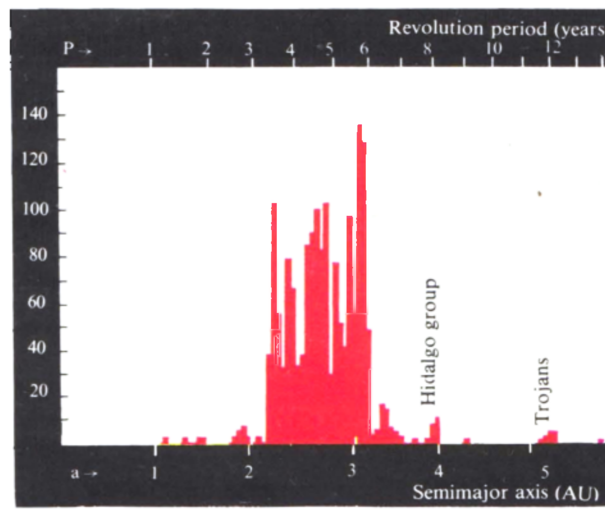
The rapid increase in the number of newly discovered asteroids at one time terrified orbit calculators – those unassuming and indefatigable workers who have taken on the job of calculating the orbits and the ephemerides of the "pocket planets". The chief worry and labour was calculating the perturbations in the motion of the asteroids. They are, especially some of them, annoyingly close to Jupiter, whose enormous mass produces large perturbations. Jupiter causes the orbits of many minor planets to change so rapidly and so much that unless the orbits are constantly corrected there is the risk that the asteroids would be lost again among the countless faint stars. There are no longer enough skilled hands, or if you wish eyes, to make the continuous observations and corrections for perturbations. The minor planets are "serviced" by two or three large astronomical institutes, among which is the Leningrad Institute of Theoretical Astronomy of the USSR Academy of Sciences. Everyone has begun to lose heart and to talk of establishing approximate orbits for the asteroids and then following only the more interesting of them – "you can't chase all of them." But then fast computers came in and made the life of astronomers easier. The Institute of Theoretical

Astronomy in Leningrad developed special methods for computing the perturbations rapidly and accurately. Based on its calculations observations of minor planets are being carried on in a number of countries.

It would be wrong, however, to think that the discovery of the asteroids has given us nothing but fruitless worry. The existence of a whole ring of asteroids in the solar system is itself very interesting and important for an understanding of the past and future of the planets. The problem of the asteroids, it turns out, is linked with the riddle of the origin of comets and the rocks (meteorites) that fall onto the ground from interplanetary space. The orbits of the minor planets and their perturbations have presented theoretical astronomers with a number of new and difficult problems many of which have been solved brilliantly and applied in other fields of science, in particular in studying the motion of electrons within the atom.

The perturbations observed in the motion of many asteroids have helped in determining the masses of the major planets. Eventually, observers were very interested in the new discoveries, and in order to find the asteroids they went on assiduously improving their instruments and techniques of observation. In particular, the necessity of searching for faint planets among equally faint stars required accurate star maps to be compiled, the uses of which are numberless. The minor planets have made it possible to establish the Earth's distance from the Sun with great accuracy. Let us take account of this, and without ridiculing the efforts of astronomers whose labours remind scoffers of the sophism concerning the omnipotency of the Creator*, we shall confine ourselves to getting acquainted with the most amazing members of this amazing family of planets.

The orbits of the asteroids, running chiefly between the orbits of Mars and Jupiter, frequently differ from those of the major planets in their large inclinations to the ecliptic and their large elongations (high eccentricity). Asteroids have inclinations up to 43° (Hidalgo) and eccentricities up to 0.65 (also Hidalgo). In recent years an especially large number of steeply inclined and extremely elongated orbits have been discovered, notably for small asteroids. In this respect the orbits of



The distribution of asteroids depending on the length of the semimajor axis of the orbit and revolution period. It can be clearly seen that the majority of asteroids are to be found between Mars (1.52 AU) and Jupiter (5.20 AU). A small group of asteroids (Trojans) have similar orbit parameters to Jupiter, and those of Hidalgo group have two-thirds the period of Jupiter. This brings out distinctly the gravitational effect of Jupiter.

the asteroids are an intermediate link between the nearly circular orbits of the major planets and the very elongated orbits of the comets. The orbits of Hidalgo and certain other asteroids are even more elongated than those of a number of comets.

Our Nearest Neighbours

Of especial interest for us are the asteroids that have perihelia closer to the Sun than Mars. The first of these, and for a long time the only one known, was Eros, discovered in 1898. When the Earth and Eros are both where their orbits are closest together they are only 22 million kilometres apart, 2.5 times less than the minimum distance between the Earth and Mars. At this time the position of Eros against the stars as observed from opposite points on Earth differs by almost one minute of arc. Knowing the diameter of the Earth and having measured the difference in Eros's apparent position in the sky, we can accurately work out the distance from Eros to the Earth in kilometres. But since Eros's orbit is known, this distance may be expressed in terms of the distance of the Earth from the Sun, and a comparison of these two quantities will

* The sophism about the omnipotence of the Creator, as is known, consists in the question: can the Creator, if he is omnipotent, make a rock so big that he could not move it?

then give us the distance of the Earth from the Sun in kilometres. The distance of the Earth from the Sun is the unit by which we measure distances in the Universe, and therefore the observations of Eros are very valuable. The year 1952 saw the completion of the processing of the great number of observations from the last approach of Eros towards the Earth in 1931 (the most favourable approaches recur every few decades). As a result, the distance from the Earth to the Sun was found to be 149,504,000 kilometres, with a possible error of 17,000 kilometres, or 0.01 per cent.

True, we now can determine the size of our unit, our astronomical "yardstick", more precisely, but Eros, too, enables us to determine it with reasonable precision.

The diameter of Eros is only 25 kilometres, and at its closest approach to Earth, when it is at perihelion, Eros shines like a seventh magnitude star, so that it is visible even in opera glasses. As it recedes from the Earth, it becomes fainter and is usually seen as an object of the eleventh or twelfth magnitude, but at aphelion, when it is beyond Mars's orbit, it is even fainter.

By a strange coincidence, Eros has attracted unusual attention for another reason – because of its extraordinary fluctuations of brightness. In 1900, during one 79-minute period it decreased to $^1/_4$ of its initial brightness (by 1.5 stellar magnitudes) right before the eyes of the astonished observer who had been watching it. In the next few hours it returned to its former brightness and then began to fade again. It was found that the variations in brightness were periodic, and that over a period of 5 hours 16 minutes it would reach two maxima and two minima. Astronomers had barely had time to get used to this when the light variations began to subside, and after several months they disappeared completely.

During its subsequent approaches to the Earth Eros would sometimes fail to vary in brightness at all, sometimes vary only slightly, and at other times vary with its earlier great amplitude. The mystery surrounding Eros deepened and caused astronomers to rack their brains over its puzzling behaviour.

Finally the idea began to gain hold that Eros is shaped like a cucumber, a sigar or a longish barrel or cask, and that in addition is covered with dark and light spots. The mutual position of the Earth and Eros changes. When the axis of rotation of this little barrel, being perpendicular to its length, is directed towards us, we see its full length all the time, and therefore the visible surface area reflecting sunlight is large and constant and at these times the brightness of Eros is both great and constant. When we are in the plane of the equator of this mis-shapen little planet it turns its long side, its bottom or its top towards us at different times. and then its brightness varies the most. Most often. however, we are only near its equatorial plane, and then we see partly its side and partly its bottom, and the brightness does not vary as much.

During its favourable approach of 1931 Eros's disk was discerned in a large telescope and changes in its shape were discovered – at times it seemed to be round and at other times elongated. Its thickness was estimated as six kilometres and its length at 22 kilometres. and it was also found that it rotates about its axis in the same direction as the major planets. In 1938 the Earth passed through the equatorial plane of Eros. and the expected large variations in brightness were indeed observed by Professor V. P. Tsesevich among other observers.

We have already mentioned that other asteroids. too, fluctuate somewhat in brightness, revealing their rubbly irregular shape and the spottiness of their surface, but of them Eros is apparently the least like a sphere.

Almost 35 years went by before another asteroid was discovered, which like Eros had its perihelion inside the orbit of Mars. It and the other asteroids, when they come very close to the Earth, seem, because of perspective, to move very rapidly among the stars like some comets and even faster, although their space velocities are not great. Therefore out of caution for some time after the discovery of a body of this type, it is called an "object", for example "Smith's object" if Smith discovered it. Only after it is definitely shown that it is really a minor planet that has been discovered does it receive a real name, and cease to be called by that indefinite and, I would say, even somewhat insulting word "object".

The object discovered in 1932 and subsequently named Amor (or Amour) proved to be an asteroid which intersects the orbit of Mars and occasionally approaches us to within 0.1 AU. It was observed again in 1940 and in 1948.

Apollo, which comes closer to the Earth than Eros or Amor, was also discovered in 1932. Its revolution

period is less than Mars's, only 1.8 years – the first case of this type to be found in the asteroid family. Apollo intersects the orbits of the Earth and Venus while its aphelion lies beyond Mars's orbit. Discovered during one of its favourable approaches, it passed the Earth only three million kilometres away (Eros's distance is seven times larger). Rapidly receding from the Earth, this dwarf planet ceased to be visible.

When we speak of an asteroid's and the Earth's orbits intersecting, this should not to be taken literally. If it were so, the Earth and an asteroid could, of course, at some time collide.

In all such cases the orbits of the asteroids are inclined to the ecliptic and just intersect the plane of the ecliptic, and not the Earth's orbit itself. Intersection of orbits themselves only occur in a plane, i. e. in the drawing, in the projection on a plane.

In a drawing the orbit of an asteroid sometimes appears to intersect the Earth's orbit, but in actual fact the asteroid is either much higher or much lower than the plane of the drawing, that is, the plane of the ecliptic. There is so much empty room in space that it is almost impossible to have a collision there.

An object discovered in 1936 also proved to be an asteroid and received the name Adonis, and an object discovered in 1937 is now listed as the asteroid Hermes. Again, the perihelia of both little planets lie inside the orbit of Venus, and the orbits are extremely elongated.

A little planet of 1949 was named Icarus for its "audacious" approach to the Sun at perihelion. At aphelion Icarus is in the normal region for asteroids, while at perihelion it goes closer to the Sun than Mercury, being only at $^1/_5$ of the distance of the Earth from the Sun and receiving 25 times as much heat as the Earth. In such scorching heat as Icarus receives when at perihelion, its surface is heated, perhaps to the point that it even begins to shine a little by its own light.

The ancient Greeks had a myth that Icarus wanted to fly and made himself wings out of feathers stuck on with wax. But he was not cautious going too close to the Sun on his wings and the Sun's heat melted the wax. Icarus fell to the ground and perished, punished for his audacity. We trust that nothing like this will happen to the asteroid Icarus, at least until we have been able to study its motion better. There is every reason to believe it won't, since this little planet is, of course, not made of wax but probably of rock.

In June 1968 Icarus again approached the Earth, an event astronomers had waited for almost 20 years. But beginning in 1965 rumours were afloat among ignorant people that this approach would cause earthquakes and floods. Nothing of the sort, of course, happened as this tiny planet on 15 June 1968 passed the Earth at about 7,000,000 kilometres, i. e. at a distance 20 times larger than the Moon's. And if even the lunar attraction only produces weak tides in the oceans and has no other impact on the Earth, how could the tiny Icarus inflict any damage? Even the lunar tidal force 16 times further away than it is now would be about 1.5 thousand times smaller, i. e. quite intangible. There have been cases when minor planets have passed even closer.

The meeting with Icarus was good, not bad, for us. According to Einstein's theory of relativity the perihelia of planets close to the Sun should slowly shift in space. To check this by the motion of Icarus would be more precise than by Mercury. From the same theory the distance of a planet to the Sun as it revolves around our luminary should vary in a somewhat different manner than according to Newton's theory of gravitation. This could also be tested. So we await new meetings with Icarus!

Adonis flew past the Earth at a distance of 1.5 million kilometres, but Hermes took the cake by tearing past us only one million kilometres away, which is merely three times the distance to the Moon. In astronomical terms it was then "just round the corner". If Hermes's orbit is not greatly altered due to perturbations (which must be great owing to the planet's close approaches to the Earth and Mars), it could one day approach us at only 500,000 kilometres, only 1.5 times the distance to the Moon.

It is extremely difficult to discover an asteroid of the Hermes type. In the first place, it can only be seen for a short time while passing near the Earth and is therefore bright enough. As it moves away it rapidly grows fainter and drops out of sight. In the second place, owing to perspective, this kind of body when near the Earth moves against the background of the stars with extraordinary speed. So on 30 October 1937 as it was rushing past the Earth, Hermes looked like a star of the eighth magnitude and was covering five degrees an hour. In twenty-four hours it crossed several constellations, and it was not surprising that many people, hav-

ing received a telegram informing them of the discovery and position of the planet, tried but were unable to photograph it. Moreover, although it should have been on several photographs, its trail could not be found; its image had moved so rapidly across the plate that it had failed to leave an impression. For observers it shot across the sky like an express train past a man who had suddenly stepped out of the woods near the tracks. Astronomers later managed to establish Hermes's orbit by using photographs that happened to come out since it was then moving slower. The difficulty in discovering such asteroids together with the fact that a number of them have been found in recent years shows us that there must actually be very many asteroids that come close to the Earth, and that they must be very small. The number of asteroids having a perihelion distance about one astronomical unit must be very much greater, because those that are moving far from the Earth shine too faintly and will only be detected in the future. The asteroids that almost intersect the Earth's orbit do not pass near the Earth on anything like once a revolution around the Sun.

By 1974 already 34 asteroids were known that at times venture far beyond the ring of the bulk of the minor planets and can on occasion approach the Earth. Broadly, they fall into two types:

1. *Apollo type*, with a perihelion of less that 1 AU, and hence penetrating inside the Earth's orbit. They are 17 in number, of which five were discovered after 1971 and only three have permanent names.

2. *Amor type*, with a perihelion from 1.00 to 1.26 AU. Also, 17 of them are known.

Each asteroidal group can be subdivided into subgroups: (1) those with a semimajor axis less than that of Mars ($a<1.52$); (2) those lying mainly between Mars's orbit and the asteroid belt; and (3) the members of the belt ($2.12<a<3.57$).

The record holder as of 1 January 1974 was Icarus: the shortest period (1.12 years) and smallest semimajor axis. It also features the largest orbital eccentricity (0.847) and the smallest perihelion of 0.187 AU. The asteroid 1973 NA has the largest orbital inclination (68°). As for close approaches, they can only be predicted for a few years because the larger planets strongly perturb asteroidal orbits, so that these asteroids may be easily "lost". With this reservation the record 0.003 AU (about 500,000 kilometres) lies as yet with Hermes (1937). The smallest asteroid, it seems, is Adonis with a diameter as small as 300 metres.

The origin of the asteroid belt, as well as the larger planets, is dealt with in the sequel to the book, but the tiny planets just discussed will also be touched upon in "Where Were Comets Born?".

Travels To and On Hermes

What would we encounter if we set out in a spaceship for the asteroid Hermes? What could we see from it, and how would we feel on it?

First of all, if we were travelling on Hermes we would be able to see the Earth in a wide variety of shapes, in particular, in the shape of a crescent or a half-circle. The Earth, as seen through a telescope, never resembles a school globe either from the Moon or from any other celestial body. When observing the Earth from Hermes we would almost always be at the disadvantage we usually are when observing the Moon from the Earth and for the same reason: the sphere is illuminated by the Sun's rays from the side. The colouration of the Earth's surface would be governed by the colouration of the ground and the vegetation, while on a school globe the colours alternate arbitrarily to denote peaks and lowlands, or to denote the territories belonging to different countries.

Almost half the Earth's surface would constantly be hidden by a white, in some places grayish, veil of clouds which at times would obscure the land for days or even weeks with a solid, monotonous cover, and at other times would move over it in broken swirling heaps. Through the gaps some details of the surface might be glimpsed, though they would be difficult to make out. Occasionally the veil of clouds would grow thin, would melt away, and then suddenly parts (only parts!) of the familiar outlines of continents and islands would appear. Something like this picture has already been observed by astronauts from spaceships. The change in colour of the continents on the Earth would probably puzzle inhabitants from another planet. Huge bright green areas would for some reason gradually turn yellow or brown, and then suddenly, after being covered by a veil of clouds emerge from it, amazing them with their unexpected blinding whiteness. In places it would seem that this blinding white colour had displaced all the green hues, but looking closely they would see that for example in Siberia and Canada, there were large dark green areas mixed with

the glittering white spots. After several months the white spots would take on a light gray colour and then become dark gray and finally would again turn to green, while the large areas of the African deserts would invariably retain their yellow hue, turning neither white nor green throughout the year.

These strange seasonal, and occasionally rapid, changes of colouration, which are caused by snow falling in winter and the grass and forests turning green in spring, would have stumped us if we were the permanent inhabitants of Hermes or the Moon. On these bodies snow and clouds are unknown, and on the Moon the alternation of the seasons is also unknown, and lacking experience we could only conjecture theoretically about what consequences the inclination of its rotational axis to its orbital plane would have for a planet with an atmosphere.

We might hold many long discussions constructing hypotheses to explain why some areas on Earth sometimes turn white, and having surmised the cause, argue about the depth of the snow cover on the ground, about whether life under this snow cover is able to make it through until spring and the warmth; can there be organic life on Earth, and aren't the seasonal changes in the colouration of the continents a sign, not of vegetation, but of seasonal chemical reactions in the soil from the effect of the changing solar illumination?

Yes, without analogous experience, a scientist living on another planet would have a hard time imagining the true conditions of life on the Earth and their diversity that depends as it does on the climatic zones.

The Earth, as seen with the naked eye from space, must look like a greenish star – this has been ascertained by studying light from the Earth reflected by the Moon. By reflecting light from the Sun onto the Moon the Earth makes visible the portion of the surface of our satellite that is not illuminated by the Sun and is, as it were, enclosed by the bright lunar crescent. The so-called Earth-shine on the portion of the Moon not illuminated by the Sun is the more noticeable the narrower the bright crescent is. It is not difficult to ascertain that in Hermes's sky at its closest approach the Earth shines about 15 times brighter than does the full Moon here on Earth. In the lunar sky full Earth is 80 times brighter than full Moon in the Earth's sky. As seen from Venus the Earth glitters six times brighter at the time of its greatest brightness than Venus does in the Earth's sky, and in Mars's sky the Earth is as brilliant as Jupiter is to us. On the other hand, from Pluto the Earth would be such a faint star scooting back and forth always near the Sun that it would be difficult even to find this fidgety midget in the glare of the Sun. It would be even more difficult, virtually impossible, to establish the existence of the Earth from the nearest star even if we had there the most powerful telescope built on Earth.*

But let us return to Hermes and from there once more look at the Earth through the telescope. We note that details of its surface structure are visible in the center of its disk.

At the edges the details of the Earth's surface are seen dimly under an overlying bluish-white veil caused by the scattering of sunlight in the Earth's atmosphere. The cause is again the air ocean that, as it were, throws a veil over the surface, a bluish haze, that softens and blurs the outlines. Recall how colour differences fade and how all shades turn bluer when we gaze at distant mountains through a thick layer of air, or look at the landscape from an airplane flying at a great altitude. Incidentally, when observed from Hermes, mountain ranges would sometimes reveal themselves by casting long shadows, but owing to the atmosphere these shadows would not be as sharp or as dark as the shadows cast by the mountains on the Moon.

One of the most outstanding details on Earth as seen through a telescope would be the glaring white spots at its poles. As the white spot at one pole grows smaller (melts) the other one grows in size and is joined by larger and larger areas that are turning white (fields and forests becoming covered with snow). The equatorial and tropical regions never turn white this way, but the former is long covered by a veil of clouds that pour down tropical rainstorms. The reflection of the Sun off the surface of the oceans is even brighter than it is off the snow.

From the Moon or from Hermes at its closest, terrestrial objects down to 500 metres in size could be seen through the best modern telescope as separate points of indeterminate shape. We could see the area occupied by Moscow and many smaller cities as a kind of spot, we would notice the veil of smoke (from chimney stacks) rising above them, but whether we would at the same time notice anything indicative of the exis-

* With the more powerful telescopes that people will build this would, of course, be possible.

tence on Earth of animals, much less reasoning beings, is questionable. Recall what I described in "Is There Life on Earth?" Thus the saying "the onlooker sees most of the game" does not always hold.

However, we began with an Earth scientist who has a good knowledge of everything that occurs on Earth setting out for Hermes. We would plunge with Hermes into the region of the asteroids. Some of them we would overtake, and others would overtake us. Thousands of asteroids would parade past us, and as they passed by, some of them would for a while outshine all the major planets. As it passes through its aphelion the sky of Hermes would be filled with myriads of wandering planetary bodies, of which only five can be seen by the unaided eye on Earth and have been known to the Earth's inhabitants since the earliest times.

In 1973 the US space station *Pioneer 10* travelled through the asteroid belt. It took it months to cross the belt. An astronomer travelling with Hermes for several years would see many wonders and solve many riddles, but unfortunately, a flight on Hermes would hardly be a comfortable one.

It should be recalled that there is no atmosphere at all on the asteroids, and there would be literally nothing to breathe there. The lack of an atmosphere would not only deprive us of breathing, but it would also deprive us of our accustomed air pressure on the outside of our bodies. Not the least of our surprises and anxieties might, perhaps, be caused by the insignificant gravity on this midget planet, only about one kilometre in diameter. The force of gravity there would be only 1/10,000 as strong as it is on Earth, that is, it would be virtually absent.

An incautious leg movement, and we would leap high above the planet, dropping back to it slowly. From a height of one metre it would take us 42 seconds to "fall" back to the planet. While "falling" from this height we would have time to drink a bottle of milk (or beer, depending on your taste), but there's still the big question of whether you'd manage to get it drunk. In such low gravity a liquid, lazily running out of a bottle, would tend to form into a ball under the action of surface tension. When struck, drops would break away from the ball and break up like mercury. But don't get the idea that you could put your lips to this "drop" of milk in the bottle and draw the liquid in. The first instant your lips touched it the liquid would run all over your face, enveloping your nose, your eyes and then your whole body. It would be possible to drink, however, by squeezing the liquid from a rubber bottle directly into the mouth. You need not, however, fear that each bite of bread will take half a minute to fall to the stomach. Food and liquid do not reach the stomach through the action of gravity, rather it gets there owing to something like spasms of the esophagus. If liquids got to the stomach solely because of gravity, the poor giraffes would never be able to drink, or else for every swallow they would have to lift up their heads with their mouths full of water.

It is not the "feeding problem" on Hermes that is frightful; most of all you must beware its mischievousness. Don't take it into your head to do any jumping for joy. One slight jump up, and you will remove yourself forever from Hermes and move away into interplanetary space. The fact is that the velocity that your muscles could impart to your body in jumping would be higher than its critical velocity and would overcome the gravitational attraction of the asteroid. In this regard it wouldn't even be safe to walk on Hermes, and in order to ensure you wouldn't part from it and the rocket that brought you to it from Earth, you might, perhaps, have to walk on your hands, or rather hang on with your hands to its very probably rough and angular surface.

If all this makes you angry, don't lose your temper and throw something at this tricky planet. Owing to the law of action and reaction, the effort would impart to you a backward push and suddenly shove you off again.

The low gravity, and hence the low weight of your body, will enable you – to the envy of Indian fakirs and dervishes – to sleep painlessly on the sharp stone razors of Hermes's surface. A clock with a pendulum one metre long would run so slow here that it would have to be discarded. A single oscillation of the pendulum, instead of lasting one second, would drag out for several minutes, and a clock keeping correct time would need an invisible one-second pendulum, since it would be only a fraction of a millimetre long.

A stay on Hermes would show us much that is surprising and interesting, but an account of it would take us too far afield. We shall limit ourselves to pointing out that at its perihelion Hermes would subject us to the scorching rays of the Sun several times stronger than arrive at the hottest places on Earth, while at

aphelion it would plunge us into protracted cold spells surpassing the Earth's. Along the way we would be pierced by cosmic rays and baked bŷ the unrelenting ultraviolet and X-rays of the Sun; we would be bombarded with prodigious velocity and force by the tiny dust particles and pebbles that fill space. Special protective measures would be needed against these, for us unusual, effects.

You could tie yourself to Hermes with a rope and pull yourself back onto it if you should accidently shove too hard against it. You could put on a special suit – a space suit – pressurized and ventilated by air from a supply of compressed air worn on the back. The Sun would warm you in the suit just about the same as on Earth, and it would be easy to keep from getting too warm at perihelion by arranging something to provide intermittent shade. If necessary electric heaters would protect you from freezing, and you could easily carry all this around with you since no load would be too much for you there. Why, a load that would weigh 10 tonnes on Earth would weigh only one kilogramme there! A man could easily hold a loaded freight car on his shoulders.

Much of what has been said here about an imaginary journey to Hermes I wrote more than thirty years ago, when space flights seemed just a daydream. Now this concerns real astronauts very much as they travel in spacecraft around the Earth and walk in space.

After this brief excursion into the realm of interplanetary travels, which have already ceased to be a fantasy, and after examining the significance that the study of these midget planets has, we shall turn below to some objects that are even more minute.

3. Visible Nothing

Heralds of Terror

An aged Russian chronicler dipped his goose quill into the ink and carefully wrote on a scroll: "In the year 7127 (1618) there was a great sign: a star appeared in the skies over Moscow. In size it was the same as the other stars, in brightness it was brighter than they were. It was above Moscow, and its tail was large. And its tail pointed to the Polish and German lands. From the star itself there ran a narrow tail, and with time it began to stretch; the tail seemed to stretch out to a thousand paces. When they saw this sign in the sky the tsar and all the people were greatly afraid. They hoped that this was a sign to the Moscow tsardom, and they were afraid of the prince who at that time had come near to Moscow. Wise men and philosophers began to interpret this star, such that the star meant not the ruin of the Moscovite state, but joy and peace. Of the star they interpreted thus: in the state over which it rests its head there will be no mutiny, but in the states over which it rests its tail there will be every kind of disorder and much bloody internecine fighting and great wars between them. Thus did the interpretation come to pass."

Here was something to think about, and the chronicler inquisitively turned over the yellowed scrolls penned by the diligent hands of his predecessors.

Under the year 1066 in the Kievan chronicle he read:

"At this time there was a sign in the west, a great star, with blood-red rays, that rose in the evening with the sunset and stayed for seven days; this manifestation was not for good, for by it there was much internecine war and the invasion of the Russian lands by the unholy, for this star was blood-red, manifesting bloodshed."

At the same time in Lykhny (the old capital of Abkhazia) the local chroniclers had inscribed on the wall of the church in the intricate Georgian script: "This occurred in the year 6669... in the reign of Bagrat, son of Georgy, in the thirty-eighth indict* in the month of April. A star appeared from whose womb there came a long ray connected to it. This continued from Holy Week until the full moon."

And at that same time in distant Normandy, at Bayeux, Queen Matilda of Flanders was weaving an image of the same monstrous star. She devoted to it a corner of the tapestry in which the Queen depicted the most noteworthy episodes of the overseas campaign of her husband William the Conqueror, Duke of Normandy. He had won a victory at Hastings in England. and was not this "amazing hairy star" – the comet – its herald? "They wonder at the star," is the inscription in Latin the Queen embroidered with the comet, which is shown above the castle attracting the attention of the people. The English King Harold is shown sitting on his throne disscussing with his astrologer the meaning of this frightful and mysterious celestial phenomenon.

The Normans, the Russians, and the Georgians had all seen the same comet, later named Halley's comet. but more of that later. Right now, however, take a look at a modern photograph of a comet and try to see in it any of what one Simon Goulard in France wrote about the comet of 1527:

"It brought such great terror that some died of fright and others took sick. Hundreds of people saw it. and to all it appeared long and blood-red. At its peak they made out a bended arm holding a heavy sword, as if trying to strike them. Above the point of the sword there shone three stars and the one that touched it outshone all the others. On both sides of the comet's rays they saw a great number of axes, knives and bloody swords amidst which a great number of severed heads with dishevelled hair and beards produced a terrible spectacle."

A Russian proverb truly says that fear has large eyes!

Goulard adds to his description: "And what did all Europe witness during the next sixty-three years if not the baneful consequences of this terrible omen?" True. indeed. Thumb through the chronology of Europe and you will scarcely find five or ten years during which there were not somewhere wars and destruction, pestilence and natural disasters, or some king or prince did not die. It was nothing more than self-conceit, however, that caused these rulers to believe that it was for them that these tailed stars would suddenly appear to foretell their personal success or death.

So in the year 837 the French King Louis the Pious. despite all his piety, was none too sure of deserving the "Kingdom of Heaven", he was also afraid of losing his earthly kingdom. He was frightened by a comet and

* An ecclesiastical unit for reckoning the years, equal to 15 years.– *Translator's note.*

summoned his astrologer, i. e. from the point of view of many theologians an "impious one", since astrologists dared to anticipate "divine providence". However, the Catholic clergy and the astrologers knew how to get along with each other. They both received many privileges so they didn't bother each other. "Go," said the King to his astrologer, "onto the palace terrace and come right back and tell me what you see, for I did not see this star last night, and you have hidden it from me and not shown it to me, but I know that it is a comet and that it heralds a change of rulers in my country. Woe to me!"

In the 17th century *The History of Miracles*, a collection widely known in the West, contained the statement: "A comet is a certain sign of unfortunate happenings. Every time there is an eclipse of the Moon, a comet, an earthquake or the transformation of water into blood, and other similar miracles, soon thereafter frightful disasters will follow: bloodshed; murders; the deaths of great monarchs, princes and nobles; insurrections; treason; lands laid waste; the destruction of empires, kingdoms and cities; famine and inflation; the plague; widespread death of people and livestock. In a word, all the troubles and misfortunes that can overtake man. And therefore no one should doubt these signs and miracles, which forewarn us that the end of the world and the day of judgement are approaching and are at hand."

In the face of such fear and confusion more than one soul, made devout out of fear, donated his property in the year of a comet to a monastery in order to escape punishment for his love of earthly joys. The epidemic of fear brought both wealthy estates and beggars' last pennies to the monasteries. While they supported the expectation that the day of judgement was near, no church or monastery ever refused a gift, whether it was large or small. Anyhow, astrologers, especially the pious ones (and there were such), always had a way out, and they were not risking much in predicting the end of the world.

Yes, the comet foretold misfortune, but if it didn't occur, why, that meant that the tears of repentance and sacrifices to churches had assuaged divine anger, that the "Lord had sheathed his sword." Prophets were given a free hand, or rather, a free tongue, by the remarkable "rule" that some of them divised. The troubles and happenings might not come about immediately after the appearance of the comet, but after forty years, or after a number of years equal to the number of months (or weeks!) during which the comet was visible.

Fortunately for them the astrologers did not have telescopes or photography, or then they, as we do today, could have discovered several comets every year, of which one would be seen by one astrologer for such and such number of months and by a different astrologer with a larger telescope for a longer time. Then try and tell which troubles to attribute to which comet.

Nor did they get by without misunderstandings, and one of them was caused by the comet of 1456, which had a long, slightly curved tail. During the middle ages many of the campaigns of plunder in the Near East were elegantly called "crusades", and Pope Calixtus III proclaimed: "This comet has the shape of a cross. It blesses our crusade, let the heathen tremble!" At the same time the followers of Muhammad joyously saw in the comet the likeness of their true weapon – the yataghan – and announced that the Moslems would be victorious. When the Pope heard of this he again studied the comet, and to his horror discovered that the comet did really resemble a yataghan. The terrified Pope renounced his prophesy and cursed the comet – fiend and helpmate of the heathen. But again the comet tricked him, and the Turks were beaten near Belgrad.

"Times change," as the Romans used to say, and now another fear of comets began to predominate in the minds of those who had chanced to hear something of the nature of comets as discovered by science. The following lines of Béranger, evoked by the new ideas are not just coincidence:

God sends a terrible comet against us,
We cannot escape our fate;
I feel the end of the world is coming;
All compasses will vanish with it.
Leave off carousing, you who have drunk too much,
This feast is not suited to the taste of many,
Better go to confession, hypocrites!
We've had enough, our world has grown old...

In Béranger's time, people frequently argued about the possibility of a comet colliding with the Earth.

Rumours based on some usually incorrectly understood phrase from a scientific report would spread persistently, being not only exaggerated, but often even distorted to the point where they were unrecog-

nizable. As an example of this we can cite the panic that gripped France in 1773. The easily excited French were expecting a collision between the Earth and a comet and the end of the world, all based just on the announcement of a forthcoming lecture by Lalande at the Academy of Sciences.

One of the contemporaries wrote in his memoirs on 6 May 1773: "At a recent public meeting of the Academy of Sciences M. Lalande was to present a paper, the most curious of all that have ever been read, but because time was short he was not able to do it. His paper contains a discussion on comets that, when they approach the Earth, can cause *perturbations* in its motion. From this there arose some uneasiness, which spread farther and wider, and encouraged by ignorance, gave birth to many fables..."

"May 9. M. Lalande's study overflowed with the curious who were consumed with impatience to learn something about his paper. The agitation reached such a point where the pious, not understanding anything about this question, had solicited from the Archbishop permission to hold a forty-hour service in order to forestall the expected deluge, and the Archbishop was ready to issue the necessary orders, but the academicians convinced him of the strangeness of such measures."

Later it developed that Lalande (as is now known, erroneously) postulated that in the future some comet in coming near the Earth might cause strong tides in its oceans, but that this would occur only very slowly, and after many years. Apropos of this, a contemporary of the panic, the great satirist Voltaire wrote caustically: "Certain Parisians have advised me that the end of the world is approaching and that the world will definitely end on May 20 this year. On this day they are expecting a comet that, sneaking up from behind, will overturn our globe and transform it into intangible dust, as a warning notice from the Academy of Sciences is alleged to state, which notice, however, was certainly never issued. There is in fact nothing more likely because Johann Bernoulli in his *Discussion of Comets* predicted that the comet of 1680 would return and carry out a terrible massacre on Earth on May 17, 1719. If Johann Bernoulli has erred in predicting the end of the world, he was only by fifty-four years and three days out. Hence it is clear, and good sense dictates, that we should expect the end of the world either on the 20th of this month, 1773, or several years later, in some month of some year. And even if the end of the world doesn't come on that date, we may nevertheless hope that if it has been postponed, why, all is still not lost..."

Incidentally, the mathematician Daniel Bernoulli, who time and again expressed some wild opinions along with his valuable ideas, declared in a work called *The System of Comets* that "if the body of a comet does not serve as the visible sign of God's anger, why, it is very likely that its tail does."

It would be a mistake to think that the superstitions of medieval astronomy are never encountered in our time. Thus it is known that Hitler paid great attention and respect to astrologers and heeded their advice, but they nonetheless failed to predict his fate.

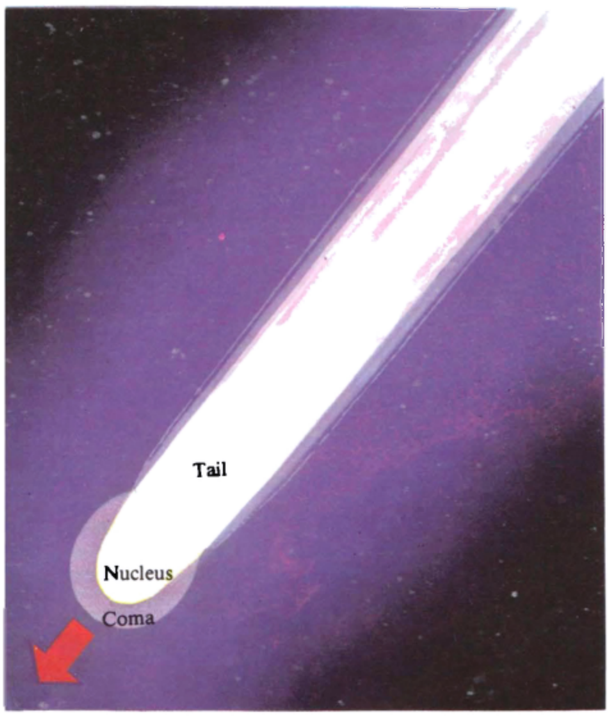
An astrologer living in the USA, J. Dumonceau, in his magazine *The Good Astrologer* (good?–for whom?) has gone so far as to cast horoscopes (predictions) not merely for newborn humans, but also for newborn acts and organizations. Dumonceau "established" that the North Atlantic Treaty was born, i. e. signed, on 4 April 1949 at 17.00 hours Washington time, i. e. at the very moment when according to his data the planet Uranus was in the centre of the magic circle, "and this planet represents both the inspiration of the supporters of the Treaty and their purpose." He finds (and in this, obviously, he is right) that the USA was the initiator of the Treaty according to the horoscope, in accordance with the position in the sky of the constellation Gemini, "the sign of the zodiac that directs its destiny". The aim of these arguments is to prove that the conclusion of the Treaty was a "necessity" dictated by the heavens...

Celestial Chameleons

From modern photographs it is easy to familiarize oneself with the variety of comet shapes and to follow the changes in shape that justify our calling comets celestial chameleons–so changeable are they.

As a rule, all large bright naked-eye comets have had tails. Small faint comets often have barely noticeable short tails visible only on photographs, and occasionally they lack tails altogether. Many comets are only visible in the telescope as small, faint, and nebu-

As a comet passes the Sun the comet's tail points away from the Sun at all times.



A schematic diagram showing various parts of a comet.

lous spots, fading out around the edges; they are called telescopic comets. But any bright comet, however, is also a telescopic comet, small and faint, when it is distant from the Sun. A comet's tail appears and grows as it approaches the Sun, and as it recedes from the Sun the tail grows smaller and disappears. When they have lost their tails, comets, like lizards, are able to grow new ones back.

The apparent size and brightness of a comet depends naturally on its distance from the Earth. A huge comet passing by far away may appear small, and vice versa. With three determinations of a comet's position in the sky made at different times it is possible to compute its orbit, and then to allow for the effect of its distance from the Earth on the appearance of the comet. Of course, for its orbit to be computed more reliably it is necessary to have, not just three, but a greater number of observations of its position.

The brightness of a comet (corrected to allow for the effect of its distance from the Earth) varies in relation to its distance from the Sun in different ways, but usually much more rapidly than by the inverse square of the distance. This fact was first established by Professor S. V. Orlov in Moscow. For example, when its distance to the Sun is reduced by half, a comet's brightness is increased ten to twenty times. This shows that a comet does not shine simply by reflected light. Otherwise a comet's brightness would vary like the brightness of the planets, i. e. according to the simple inverse square of the distance, and at half the distance from the Sun the brightness would increase by a factor of only four. The laws of the variation of comet brightness have been studied in greater detail by S. K. Vsekhsvyatsky and B. Yu. Levin.

A comet's tail, as is well known, is always pointed away from the Sun, and when a comet is receding from the Sun its tail moves in front of the comet – surely the only such instance in nature among things having tails.

A comet consists of several components that differ greatly in their nature. Misunderstandings often arise, therefore, if people discuss a particular property of a comet without indicating exactly which part of it they are talking about.

In a comet one should distinguish the *nucleus* (more precisely, the *apparent nucleus*), the *head* (also called the *coma* if the comet has no tail) and the *tail*. The head, or coma, is the brightest part of the comet, brighter in the centre, in which something resembling a small star, often a nebulous one, is usually visible. This is the apparent nucleus of the comet. It alone, perhaps, is a solid body, but it is more likely that even it consists of individual hard pieces.

Cometary nuclei are very small, and it is difficult even to measure them. In 1910 for example, Halley's comet passed directly between the Earth and the Sun. If its solid, opaque nucleus had been more than 50 kilometres in diameter it would have been visible as a black dot against the radiant background of the Sun's disk. Nevertheless nothing like this was noticed, not even the slightest shadow passed across the Sun. In 1927 the Comet Pons-Winnecke approached very close to the Earth. Not even the tiniest disk at its nucleus was noticed in powerful telescopes. Whence it follows that it was less than two kilometres in diameter. From estimates of its brightness, assuming that it is a solid body and reflects sunlight to the same degree as the Moon's surface, it was possible to conclude that its diameter is only 400 metres. It is more likely, however, that the nucleus consists of not one but of many chunks, even smaller in size and separated

from one another. This conclusion is supported by many other facts with which we shall become acquainted in later chapters.

Sometimes a star-like cometary nucleus is surrounded by a fairly sharply outlined bright haze, which some observers also include in the concept of the nucleus. Misunderstandings sometimes also arise from this.

The nucleus of a telescopic comet, and of faint comets in general, is always surrounded by a large nebulous mass, quite blurred around the edges. It is more or less round in shape and brighter towards the nucleus, but often as it approaches the Sun it becomes elongated. Then its elongation is directed along a line joining the comet's nucleus and the Sun. Sometimes from this nebulous mass or coma a thin, bright ray, often several rays, grows out in the opposite direction to the Sun, giving the comet the appearance of a green onion. In the brighter comets, this thin "onion" tail develops into a long broad tail as they approach the Sun and then the coma comes to be called the head.

The front part of the head, or the envelope of the cometary nucleus as it is also called, has the shape of a paraboloid. If we rotate a parabola about its axis the surface it generates will be a paraboloid. There have been cases where several envelopes have formed, nested within each other like the wooden balls that children take apart.

The year 1957 presented us with two bright comets with remarkable tails. One of them was discovered by Arend and Roland in Belgium, and the other by Mrkos in Czechoslovakia.

When a comet moves away from the Sun these phenomena occur in reverse order, i. e. the tail becomes shorter and not as bright, then there remains only an elongated coma, and finally the comet turns into a tiny nebulous spot with a nucleus, or even without one.

The tail appears, develops and changes differently in different comets, and even in the same comet there is no symmetry with respect to the moment of its passage through perihelion. It happens that on some days the tail will suddenly grow dim and then grow stronger again. The overall brightness of a comet also occasionally shows irregular fluctuations. Several comets have been observed to have, usually temporarily, two or even three tails at the same time, although an inexperienced observer could always mistake the straight or slightly curved rays that form a single tail for separate tails. Something of the sort was discovered in 1944 by the Soviet astronomer Orlov, from studying drawings of the comet of 1744, which, according to its contemporaries, had six tails.

It is often observed that bright clouds are ejected from the nuclei of large comets now and then, sometimes at intervals of only a few hours. They move off gradually into the tail and with the passage of time seem to fade away.

The combination of these observations, especially in conjunction with the changes in the spectra of comets (to be discussed below) gives us a picture of comets as extremely capricious and changeable creatures. The changeability of these celestial chameleons complicates their study, but at the same time it permits probing deeper into the mysteries of their structure and development. But before going into greater detail concerning the physical nature of these hairy celestial wanderers we shall devote some attention to their motion.

Halley's Discovery

Newton's good friend, Edmond Halley (1656-1742), had a soft spot for comets. Halley's great teacher, having discovered the law of universal gravitation, proved that according to this law two bodies may move around their common centre of gravity only along one of the types of conic sections: an ellipse, a parabola, or a hyperbola. Newton proved that since the mutual attraction of the planets was small in comparison with the mighty pull of the Sun, each of them trace a nearly regular ellipse around the Sun.

The elliptical nature of planetary motion was known since the time of Kepler. Motion along an ellipse proved to be one of the frequent forms of motion possible under the effect of gravity. Newton was interested in whether the other forms of motion, which he had deduced theoretically, were present in nature, and he was also interested in how comets moved in space. Were they, like the planets, also subject to the Sun's attraction and to the laws of motion resulting from it?

Newton invented a method of determining the orbit of a comet from a few observations of its apparent position in the sky among the stars at different times. Developing this idea he applied it to observations of

the comet that appeared in 1680 and calculated its orbit. It proved to be a parabola whose focus coincided with the Sun. It was thus proved that comets, like the planets, are subject to the Sun's attraction and move in the solar system under its influence, tracing out a curve of a different shape than the one traced by the planets, but one that is also predicted by the theory of gravitation. But Newton did not think that *all* comets moved along parabolas, i. e. he did not think that after coming out of infinity and going around the Sun they all went back off into the infinite yonder, never to be seen again. Of the comet of 1680, which he himself observed, Newton noted:

"Observations and calculations of the path agreed as well as the calculations of the planets' paths usually agree with observations. The revolution periods of comets cannot be determined from similar calculations. They may only be found by comparing the paths of comets that have appeared at different times. If it proves that some of them, appearing after equal intervals of time, describe identical curves, then it will have to be concluded that these are successive appearances of the same comet. Then we shall determine the nature of the orbit from the revolution period itself and find then an elliptical orbit. In order to do this it is necessary to calculate the paths of many comets, assuming their orbits to be parabolic, and then compare them with each other."

Newton was very busy as it was, and so this laborious task was undertaken by Halley. He began by improving Newton's method. Then from various books, including some very ancient ones, Halley collected observational evidence of the positions and motion of the comets over a period of three centuries, from 1337 to 1698. When he had finished this work Halley wrote:

"Having collected observations of comets from everywhere, I drew up a table – the fruit of extensive and exhausting labour – that was small, but not without use to astronomers... The reader of astronomical works should take notice that the data presented here I have obtained as the result of the most precise observations and have published them only after many years of conscientious study, after having done my best."

Recalling what Newton had pointed out, Halley then began to compare the orbits of comets covered by the table, and here is the conclusion he drew:

"A great deal compels me to believe that the comet of 1531, which was observed by Apian, was identical with the comet of 1607, described by Kepler and Longomontan, as well as with the one I myself observed in 1682. All elements agree precisely, and only the disparity of periods, of which the first is equal to 76 years and two months and the second is 74 years and 10½ months would seem to contradict this, but the difference between them is not so great that it cannot be attributed to some physical cause. We know, for example, that Saturn's motion is so strongly disturbed by the attraction of the other planets, notably Jupiter, that the revolution period of Saturn is only known to us within an accuracy of a few days. Just imagine how much more subject to these effects must be a comet that drifts four times as far from the Sun as Saturn and whose speed, if increased just a little, might transform its elliptical orbit into a parabolic orbit. By such reasons do I account for the disparity in the periods of the comet, and I have therefore confidently decided to predict the return of this same comet in 1758. If it does return there will no longer be any reason to doubt that other comets too should return... but many centuries will pass before we learn how many such bodies there are revolving around their common centre, the Sun..."

Posterity named this comet after Halley. Later on several other comets were named after scientists who had made especially thorough studies of their motions. So the comet long known as Comet Encke subsequently came to be called Comet Encke-Backlund after the director of the Pulkovo Observatory, who studied the peculiarities of its motion.

Halley's comet did not betray the hopes of the man whose name it bears. It did return. But by the time that it, albeit not of its own volition, was ready to return many events had occurred on Earth. The science of the heavens had forged far ahead in this period of time. It had become possible to allow for the effects of the perturbations produced by the planets on the motion of Halley's comet. This made it possible to predict the reappearance of the comet more accurately. This large and important work was undertaken by the French mathematician Alexis Clairault (1713-1765).

Few people know what the connection is between Halley's comet and beautiful light-pink or blue flower known as hydrangea (hortensia in French). Its homeland is Japan, and it was first brought to France at the time of the return of Halley's comet. The Paris Academy of Sciences named this new (to Europe) flower in

honour of the woman who had been Clairault's faithful helpmate in his calculations, Hortense Lepaute. One of the first women scientists, she may on that occasion have recalled the fate of her distant precursor, the first woman astronomer, Hypatin. Many centuries ago (4th century) in sweltering Egypt, at Alexandria, Hypatin studied the courses of the celestial bodies and was torn to pieces for "witchcraft" by the same benighted and brutal mob, led by Christian clergy, that burned down the greatest storehouse of ancient learning, the famous library of Alexandria.

Clairault and Lepaute indicated more accurately the time at which Halley's comet was to pass through its perihelion in the middle of April, 1759, and computed its apparent path across the sky. As the comet approached the Earth and perihelion, scientists began to watch for it, but they were all beaten by J. Palitzsch, a peasant living in the vicinity of Dresden, who discovered it in December 1758. From later observations it was ascertained that Halley's comet had passed perihelion 31 days sooner than had been predicted.

At its next appearance in 1835 Halley's comet passed through its perihelion only 9 days later than had been expected according to new and more accurate calculations.

In 1910 during its last observed appearance it was only three days late for its perihelion, so well had the allowance for perturbations and the accuracy of comet observations been perfected. What will be the error in the calculations for the next appearance of the comet? This will be in about 1986, and the author hopes that all of you, dear readers, will see it, but alas, he himself does not count on being alive then.

As we see, Pushkin's words "... like an illegitimate comet in the calculated circle of the planets" can only be true in relation to comets observed for the first time.

During its last return Halley's comet was first seen on 11 September 1909 at the place predicted for that day, when its distance from the Earth and the Sun was equal to about twice Mars's distance from the Sun. After the comet passed through its perihelion on 20 April 1910 it disappeared from view (or rather, from the glass eye of the astrograph, which in collaboration with the photographic plate followed it for a month longer than it could be followed visually in the telescope) on 1 July 1911. At about this time on its path away from the Sun it had already crossed the orbit of Jupiter.

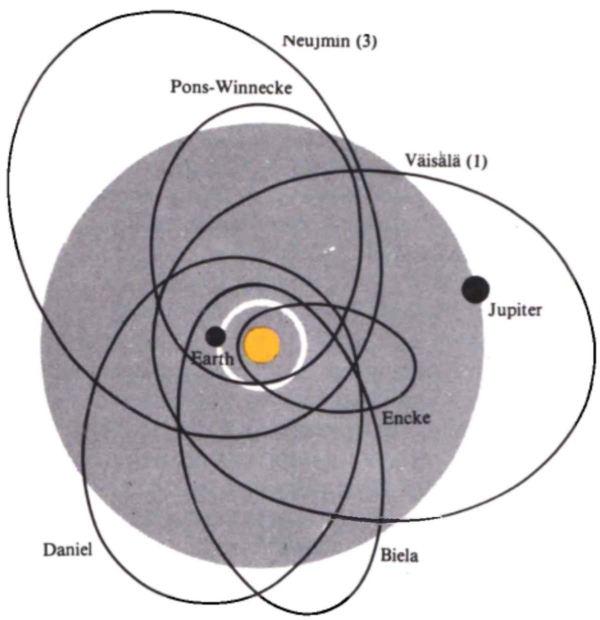
This example gives an idea of how long a bright comet remains visible and at what distances from the Earth and the Sun it can be followed.

From ancient annals and chronicles composed in ancient Russia and in other countries many previous appearances of Halley's comet have been discovered, beginning in the year 240 B.C., when it was seen and noted down by the Chinese. During its appearance in 1066 it had its first, albeit very ugly, portrait made. I mean the embroidery of Queen Matilda of Flanders mentioned above, which is the first attempt we know of to depict a comet. If you are familiar with history, try to recall to what events on Earth Halley's comet has been a witness at the times of its various approaches to the Earth. However, of the changes that have transpired on Earth between any two successive returns of Halley's comet none have been as far-reaching as those that have occurred on Earth since its last visit to us in 1910.

Short-Period Comets

The short-period comets, i. e. those moving in ellipses and having short periods of revolution, are the most thoroughly studied. Of them Comet Encke-Backlund has the shortest period (3.3 years). Thirty-nine approaches of this comet have been observed. Comet Wilson-Harrington (1949 IV) had a still shorter period of revolution – 2.3 years but it was only observed in 1949, and has not been seen since.

The periods of other comets of this type are under eight years, and their aphelia, in the majority of cases, are located near the orbit of Jupiter, which is why they are often called comets of the group, or family, of Jupiter. The family includes about 50 comets. They are all further distinguished by the fact that their orbits are near to the plane of the ecliptic (their inclinations being less than 30°) and they move in the same direction as the planets. Almost all of them are visible only in the telescope, and they are tailless. The elongation of their orbits is comparatively slight, and in this respect Comet Schwassmann-Wachmann, discovered in 1927, has no equal. Its orbit lies entirely between the orbits of Jupiter and Saturn, and from the smallness of its eccentricity it is more like the orbits of the planets than the orbits of many comets.



The orbits of some of the short-period comets.

The discovery of a comet with such an orbit is of exceptional interest, since it is a sort of a bridge between the asteroids (and planets in general) and periodic comets (and comets in general). This bridge exists not only in the sense of characteristics of motion. The very small coma – small in terms of brightness and size – of this comet invites attention. If it were just a tiny bit fainter, it would be impossible to tell Comet Schwassmann-Wachmann from an asteroid. It cannot, of course, be called comet-asteroid hybrid because we are not dealing with biology, but it does indicate to us that there is a close, albeit not yet quite clear, kinship between the two families.

In order not to have to come back to Comet Schwassmann-Wachmann later on, we shall cite some extremely interesting data about it here.

Its orbital period is somewhat over 16 years, and since the time of its discovery it has been visible and observed every year – the first and so far the only such case in the history of observing comets. Also striking are its sudden and extremely strong variations in brightness. Thus, for example, on 26 November 1927, the comet was equal in brightness to a star of the 14th magnitude, and five days later it was suddenly only $^1/_6$th as bright. In March 1935, the comet was of the 18th magnitude and could only be observed photographically, not being visible to the eye in even the most powerful telescope. In April it became 15 times brighter, and by 4 May 7 times brighter still, so that over a two-month period its brightness (later falling off again) increased a hundred fold. At the same time its distance from the Earth and the Sun scarcely varied. Evidently, some kind of sudden explosions occur within the comet, either due to internal or external causes, perhaps, from a collision with some sort of small bodies (meteorites). Here at least, the Sun has nothing to do with it, although the luminescence of comets is caused by solar radiation.

Still more striking is the orbit of Comet Oterma, discovered in 1943. Its orbit, like the orbit of Comet Schwassmann-Wachmann, is almost circular, and it *lies entirely between the orbits of Jupiter and Mars.* It does not differ in the least from the orbits of typical asteroids! It is only distinguished from a minor planet by the presence of a nebulous envelope. None of the orbits of the known asteroids coincides precisely with the orbit of Comet Oterma. The comet is very faint, and it does not display any such variations as Comet Schwassmann-Wachmann.

Among short-period comets, besides those that form the Jupiter family, a small number are known to make up the families of other major planets. The comet family of Saturn consists of six members, the family of Uranus has four members with orbital periods of 30-40 years, and there are nine comets in Neptune's family, among them Halley's comet, the brightest of the periodic comets.

The aphelia of the comets of each family are grouped near the orbit of their planet, which is the characteristic by which they are referred to a particular family. With the exception of the comets of Jupiter's family, however, the real connection of the other planets with their comet families is less evident.

The number of known periodic comets is growing all the time since they are being discovered continually. As time goes on, fainter and fainter comets are being discovered, since the brighter comets have, naturally, been discovered earlier.

Up to 1925, two or more returns had been observed for 25 of the 50 comets with periods under 100 years, the others having been observed once. There are many reasons for this: for example, recentness of discovery and length of period, the poor positions for observa-

tion during subsequent returns, and finally, the disappearance of the comets. But we shall talk about all this later.

About a score of comets have orbital periods of between 100 and 1000 years, and there are 30 comets with periods of between 1000 and 10,000 years. The longer the period, the less precisely it is known. With a period of several years the error in determining it may amount to one or two weeks, but with periods of the order of a thousand years the error increases to centuries. Thus it happens that astronomers do not always recognize their old friends right away. Their clothing and physiognomy are so changeable, as we have already said, that they cannot be recognized by their outward appearance, but only from the paths that they follow. Not only do comets lack the sort of distinct "physiognomies" the major planets possess, but even their path is not a beaten track around the solar system. Often it varies greatly owing to perturbations, and this has to be taken into account, but even without the perturbations, as a result of the poor determination of long periods we do not know precisely when to expect the return of a given comet. It happens that a comet is taken for a stranger, and only later is it ascertained after studying its motion that it is an old acquaintance that has taken a new path.

Members of the Family or Aliens?

In order to learn how, when and where these hairy stars – the comets – are created it is first necessary to know if they are members of the solar family or if they are aliens who have come to us from out of interstellar space. This question may seem superfluous to you in relation to the short-period comets just described, but it may be that in the unimaginably vast family of comets that, to use Kepler's expression, are scurrying about the solar system as numerous as fish in the sea there are comets of both kinds. It is possible that the periodic comets are adopted children of the Sun or that, on the other hand, nonperiodic comets are children of the Sun who are trying to go astray.

Alas, it has turned out that answering this question is not easy, although many generations of scientists have laboured over it. There are a number of difficulties that we have not yet succeeded in surmounting. Strictly speaking, two problems need to be solved. First: are there in general at the present time comets that are *actually* moving in open orbits – parabolas or hyperbolas? Second: can planetary perturbations change the elliptical orbit of a comet into a hyperbolic orbit, or vice versa? True parabolic orbits are in themselves quite unlikely, since to move in them a comet must have a strictly constant velocity. The minutest difference in its speed from this fixed value is all that is needed to make the orbit either elliptical or hyperbolic. True, both will be very close to a parabolic orbit. In other words, the eccentricity of an orbit close to a parabola may differ from the eccentricity of a parabola (equal to unity) by a triflingly small fraction, say, by one thousandth or even less.

If we take a cursory look at the list of the elements of comet orbits, we find that, for example, out of 111 comets discovered since 1900 (which have been observed more accurately and whose orbits have therefore been calculated more accurately), the orbits of 48 of them cannot be distinguished from parabolas. Of the remaining 63 there are 33 with eccentricities greater than, and 30 with eccentricities less than, 0.990. The first 33 orbits are close to parabolic, and the eccentricities of 15 of them are just slightly greater than one. Can we be sure that this is actually so?

We only observe comets in the sector of their path where they are close to the Sun and the Earth. If the orbit is large this sector is extremely small in comparison with the entire length of the orbit. Within the limits of accuracy of our observations, in such a small sector, we are unable to distinguish with any assurance, or even to distinguish at all, between a hyperbola with an eccentricity just a shade over one and an ellipse, extremely elongated, with an eccentricity just a shade less than one. They almost coincide. In other words, we may pass either a hyperbola or an ellipse with a long revolution period through the points that represent the observed positions of a comet in space, allowing for possible inaccuracies in their determination.

Since no comets have been discovered that have strongly hyperbolic orbits with eccentricites noticeably greater than one, we may assume that there actually aren't any. Another consideration favouring this is that the type of orbit a body has depends on its velocity at a given distance from the attracting centre. With respect to the nearest stars the solar system is moving at a speed of 20 kilometres per second. This speed, added to the speed of comets that are rushing towards

us out of interstellar space, would cause them to have decidedly hyperbolic orbits, and these sorts of orbits are not observed at all.

Careful calculations by Strömgren and his colleagues in Copenhagen have shown that by making allowances for the perturbations caused by Jupiter and Saturn on 16 comets. The orbits of 15 of them were elliptical before they had approached Jupiter and Saturn and while they were as far from the Sun as Neptune. Only when one of these comets had come close to Jupiter or Saturn did the gravitational attraction of these two planets accelerate it and make its orbit parabolic or slightly hyperbolic. Only after this did we discover the comet as such, for until then it had been too distant from us and had not been visible. Thus these comets had been members of the solar system until Jupiter hurled them out of it by forcing them to move in hyperbolas. Unfortunately, we are not yet completely certain that all nonperiodic comets originate in this way.

Still, there is no denying that some comets, if only extremely rarely, come to us from other stars. After all, if comets are thrown out of the solar system to be picked up by other stars, why cannot our solar system accept newcomers from outside?

The short-period comets, of which the great majority are connected with Jupiter, could in theory have been formed (owing to perturbations) from comets that previously had long periods, and such comets are the majority.

The result of Jupiter's disturbing force depends on the speed at which a comet is approaching Jupiter. Both the magnitude and the direction of the speed are important here. If we assume that the initial motion of the comet before it drew close to Jupiter was along a parabola, the result of the approach can be presented in the form of the table.

Direction of motion before approaching Jupiter		New orbit after approach
Same as Jupiter	Large perihelion distance	Ellipse, direct motion
Opposite to Jupiter		Hyperbola, direct motion
Towards the Sun	Small perihelion distance	Ellipse, retrograde motion
Away from the Sun		Hyperbola, direct motion

Comets that have gone over to hyperbolic orbits are hurled out of the solar system and lost. Comets that have switched to moving in an ellipse with direct motion, i. e. in the same direction as the planets, will approach Jupiter again and again, and each time their periods will be shortened a little more until they finally become part of Jupiter's family of comets.

Comets that have acquired retrograde motion in an ellipse at their next encounter with Jupiter will pass by "at high speed", close to the hyperbolic, and in most cases they will be thrown out of the solar system, too. This process of cometary "natural selection", operating for millions of years, has formed the family of short-period comets having direct motion and low inclination to the plane of the orbits of Jupiter and the other planets.

The rate at which Jupiter has acquired its family of comets and is attracting new members depends on the nature of their motion prior to their first critical encounters with Jupiter. If they were moving in parabolas at various angles to the orbits of the planets, Jupiter will have to change their motion greatly and will only capture the few that will pass very near.

According to some estimates, only 126 comets out of a thousand million such comets will acquire periods under six years, 839 comets will have periods less than 12 years, and 2,670 comets will have periods less than 24 years. Three quarters of the 839 comets with periods under 12 years will have direct motions and low orbital inclinations. Unfortunately, all these estimates depend on the initial, and unknown, distribution of comets' orbits. Differences in this distribution lead to substantial changes in the subsequent fate of the comets. Most probably, the great majority of comets have originated within the solar system, and stay on as its members, but most of them have revolution periods of thousands of years. Comets with almost parabolic orbits with periods of tens of thousands of years are still tied to the Sun fairly firmly, providing Jovian attraction does not interfere.

Discovery of Comets

Many codes have been invented by men for various purposes, and there are great experts in breaking them. But even Sherlock Holmes would not be able to decipher the telegram given below.

"Comet Jones 04107 October 18490 10073 22513 81101 20153 20056 76503 Smith."

Knowing the code, an astronomer easily deciphers this telegram as follows:

"Comet eleventh magnitude with nucleus discovered by Jones 4 October (19..) at 18 h 49.0 m UT, apparent position of the comet was fixed at right ascention 10 h 07 m 31.1 s, declination +25° 13′01″, moving east 1 m 53 s, north 0°56′ each day. Message from Smith." It is assumed that the experts who receive the telegram will know that Smith works at Greenwich.

We shall not weary you with the details of the encipherment of a similar fictitious telegram cited merely as an illustration because you are not going to have to be deciphering them. Note, however, that Jones at Greenwich hasn't discovered a comet yet. We wish to say, however, that both Jones and Smith at Greenwich or any other place, just like any other reader, could discover a comet. Providing it is not a previously known periodic comet on its expected return and bearing its own name, it is named after the person who first notices and reports it.

It happens that some people have discovered several comets, for example G. N. Neuymin. These comets do not confuse people because as a rule each of Neuymin's comets receives in addition a number designating the year in which it passed perihelion and a Roman numeral designating the order in which the comet came to perihelion among the comets discovered that year. For instance Comet Neuymin 1916 II is a comet discovered by Neuymin in 1916 and the second in order to pass through its perihelion. Until the dates of their perihelion are established, all the comets discovered over the last few years have been designated by the year and a Latin letter shows the order they were discovered during the appropriate year, for example, comet 1908a, 1908b, and so on.

Professional astronomers, especially in our day, generally discover comets either accidentally or whilst searching for a periodic comet whose return is expected. Faint comets are chiefly discovered on photographs, but in any case in order to be certain that it is actually a comet that has been discovered and in order not to lose it later on in the event of a spell of bad weather, it is essential to establish the direction and speed of the comet's motion against the background of the fixed stars.

The history of science holds the names of a good many comet hunters who sometimes made their discoveries using the most modest telescope or even field glasses.

Pons (early in the 19th century in France and Italy) had the good fortune to have discovered the most comets – 33 in all – and the American amateur astronomer Brooks (late in the 19th century) found 25 comets.

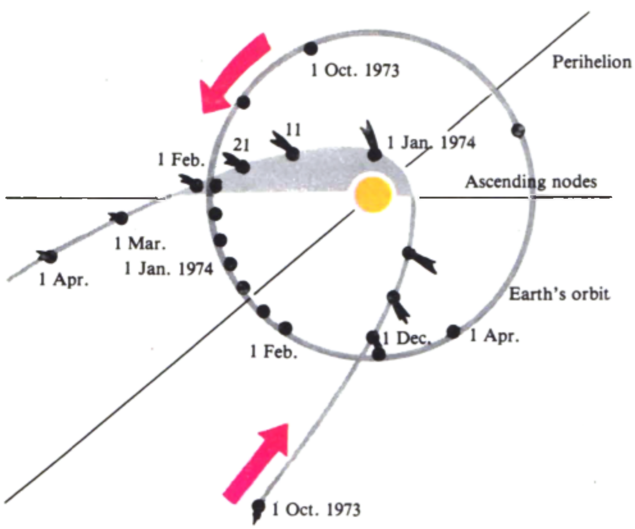
In order to discover a comet it is necessary to examine the sky carefully on moonless nights and take note of every nebulous little spot. Having discovered such a spot and having ascertained from a star map that it is not a nebula, the would-be-discoverer must still make certain that the comet is moving against the stars, and most importantly, determine its coordinates, for which in the simplest case it is only necessary to mark the position of the comet on a good star map having a coordinate grid.

This method was successfully used in Czechoslovakia by Mrkos and his wife Pajdušakova, who have discovered a number of very interesting comets.

The use of photography has made search for comets easier and has made it possible to discover faint comets, but in our day comets are still very frequently discovered visually by amateur astronomers. Special medals are awarded for the discovery of comets. Many of the comets discovered in recent decades bear the names of Soviet astronomers: Byelyavsky, Neuymin, G. A. Shain and P. F. Shain (of Simeiz Observatory), Zlatinsky (of Jelgava), Dubyago (of Kazan), Tevzadze (of Abastumani in Georgia), Kozik, Akhmarov and Yurlov, Bakharev, and others.

Some comets are discovered at a great distance from the Earth and the Sun as faint nebulous little spots, others appear suddenly with bright tails, emerging out of the Sun's rays if until then the comet had not been visible to us because it was in front of the Sun or behind it. Discovered in this way were Comet 1910a (by workers at the South African diamond mines), Comet Gregg-Skjellerup 1927 and Comets 1947n and 1948c. At the time of discovery they were all near the Sun and visible in the daytime in the bright sky.

The discovery of a comet is an exciting experience for an astronomer also because in the beginning it is unclear what will happen to the comet in the future. A newly discovered comet is a sort of a "pig in a poke". Perhaps it will remain telescopic, perhaps it will devel-



The orbit of Comet Kohoutek.

op a long bright tail and frighten ignorant people, perhaps it will pass quite near the Earth or turn out to be a very short-period comet – new or already known, or known but lost.

All of this is only ascertained after at least three accurate observations of the comet's position among the stars from which it is possible to work out its future path across the sky so as not to lose it in the event of cloudy weather. Having determined the comet's direction of motion it is necessary immediately to telegraph all the data via the nearest observatory to all the observatories of the world.

A comet discovered by observers in the northern hemisphere may then pass rapidly into the southern skies, where people in Africa and Australia can follow it, which they can only do poorly since there are too few observatories there.

An experienced mathematician can determine an approximate orbit for a new comet from three observations in five or six hours using desk calculators. A precise orbit based on many observations considered together is computed several years later, and the calculation itself took about a year before the advent of computers.

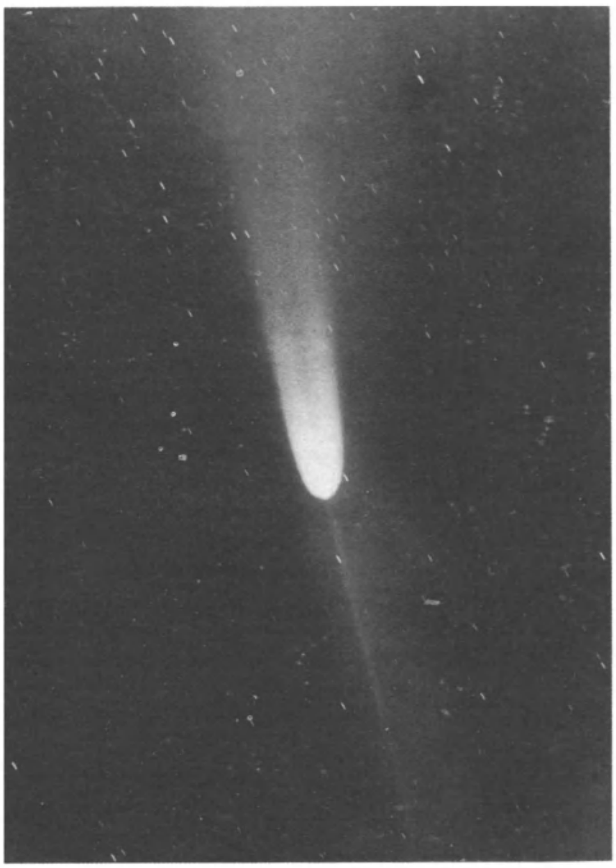
Up to 1950 about 1000 comets had been registered, of which 400 had been discovered before the invention of the telescope, i.e. were very bright. Of the 1000 comets, orbits have been determined for just over 600. The years 1948 and 1967 were the richest for comets. So in 1967, 14 comets were discovered, while in 1948 there were 10, but in addition to these 14 more comets were observed that had been discovered previously, so that during this year a total of 24 comets were observed. However, comets visible even for a short period without a telescope are a rarity. Comets having a small tail and shining like stars of the fourth or fifth magnitude were visible in the USSR in 1936 and 1943, but bright comets with long tails observed for a long period were only visible in 1910 (Halley's comet), 1957 (Comet Mrkos) and in 1973-1974 (Comet Kohoutek).

If you have the chance to see such a comet, don't miss it!

Lost Comets

Comets are not only discovered, they are also lost. Some are lost by us, while others disappear themselves, or rather, they drop completely out of sight. If we fail to obtain sufficiently accurate determinations of a new comet's position because of cloudy weather during its approach to the Sun or for other reasons, the computed orbit will be inaccurate. The subsequent actual path of the comet will deviate greatly from the calculated path, and when the faint comet is again in a position where it can be observed, it is not found in the place it ought to be.

It sometimes happens that a periodic comet after encountering Jupiter suffers such a sizeable change of orbit that it is subsequently not seen, for example, it always passes too far from the Sun and the Earth. Comet Pons-Winnecke, with a period of revolution of about six years, was discovered in 1819 and was not observed again for 40 years until it was rediscovered in 1858. Since then it has been observed almost every five years. Over the last century the inclination of its orbit has doubled (from 10 to 20°) owing to perturbations, while its perihelion, which at first was almost the same distance as Venus from the Sun, has moved out beyond the Earth's orbit. If this comet had not been observed many times and if the perturbations in its motion had not been carefully studied, it would have been difficult to have guessed that the comets of 1819 and 1933 were the same. We would have thought that they were two different comets and the periodic comet of 1819 would have remained on the roll of the missing. Allowing for perturbations, especially over a


Comet Arend-Roland of 1956 showing the antitail.


Comet Mrkos of 1957.

long period, is a difficult and laborious task. It is necessary from time to time to have verifying observations, so to speak, of the comet's position.

If after several revolutions we have failed to follow the comet and the Jovian perturbations are too great, then the predicted place in the sky where it should appear at its next approach to the Earth will be fairly uncertain. In such a case finding a faint comet lost among the countless faint stars would be as hard as finding a needle in a haystack.

In the cases described above the comets have only become lost because we have lost their paths. There are cases, however, when comets do actually vanish and cease to be observed even though their paths are well known. In many cases this probably occurs owing to a rapid weakening of their brightness, which usually comes suddenly.

An example of this is the surprising behaviour of Comet Ensor, 1926 III. The comet was of the eighth magnitude when it was discovered two months before it reached perihelion, and then because of its approaching the Earth and the Sun it should have become easily observable with the unaided eye. It would have been expected that it would grow a "decent" tail. Instead it spread out, became diffuse, faded rapidly and soon vanished. In 1913 the periodic Comet Westphal behaved the same way. Over a period of a month and a half its brightness, counter to expectation, faded by ten magnitudes.

Other comets split up into pieces before vanishing. Such was the fate of a comet discovered by the Czech amateur astronomer Biela, which had a period of about seven years. It was seen in 1772, and then on its returns in 1815, 1826, 1832. In 1846 before the eyes of

astonished observers it divided into two comets almost equal in size, each of which was fainter than the comet from which they came. Over the years the distance between them increased. Initially, the satellite comet was considerably fainter than the other one and had the same nebulous appearance. The tails of both comets were parallel and the comets almost merged with each other. After three months the satellite became brighter than the main comet and lagged behind by half the apparent diameter of the Moon. Both comets reappeared in 1852 moving in single file along the old orbit, but now far apart, with one comet being brighter and the other fainter. Since that time neither has been seen again, although they subsequently let us hear from them, how we shall learn in the next chapter.

In 1916 Comet Taylor split up in the same way as it was passing close to the Sun and has not been observed since, although the comet was not very faint and each of its fragments even had its own small tail.

According to Barnard's observations the periodic Comet Brooks was accompanied in 1889 by four travelling companions. All of them had tails, but at the next appearance only the main comet was observed; its travelling companions, having undoubtedly detached themselves from it, had, as it were, melted away.

There is no doubt that the tails and comas of comets are formed from matter that is being expelled from the comet's nucleus. The nucleus consists of hard blocks and is something like a compact swarm of rocks, as was finally confirmed by the author of this book in 1945. The tidal action of the planets and the Sun is due to their attracting the strongest parts of the nucleus closest to them. As a result the swarm of rocks is, as it were, pulled apart more and more, and eventually the nucleus of the comet is completely dispersed.

To sum up, the rolls of the numberless army of comets contain not only captured and missing comets but also those killed in action.

Visible Nothing

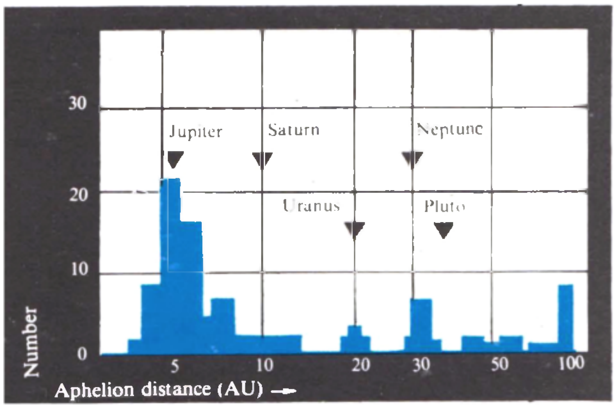
"Visible nothing" is what the French scientist Babinet called comets. By this he meant that while the outward appearance of comets is sometimes awesome and their dimensions enormously large, there is almost no matter in them. Some of these luminaries are so bright they are visible in the daytime, and their tails stretch far out, sometimes covering half the sky. Of course, other things being equal, the brightness and length of the tail depends on the comet's distance from the Earth, and when its distance is known it is always easy to calculate the comet's dimensions. A coma, or comet head, whose diameter is less than the Earth's diameter is very rarely encountered. Usually it is from 50 to 250 thousand kilometres in diameter, on average about ten times larger than the Earth. The head of Comet 1811 exceeded even the Sun in size, and the nebulosity that accompanied Comet Holmes 1892 at one time was twice the Sun's diameter. The above data on cometary dimensions were only based on visual observations of their angular diameter.

Photographs are able to bring out the fainter external portions of comets. The present author established that the head of even the modest comet 1943 I, which was barely visible by unaided eye against a dark sky, had a diameter of no less than 2,000,000 kilometres, i.e. it was about 1.5 times larger than the Sun. The size of the head of each comet varies irregularly and has a tendency to decrease as the comet approaches the Sun. For example, the head of Halley's comet, when in 1909 it was twice as far from the Sun as Mars, was twice the size of the Earth. By the time it had approached the orbit of Mars it had grown 15 times, but by perihelion it had decreased by half. When it had doubled its distance from the Sun its head again increased by a factor of 1.5, but when the comet again approached Jupiter's orbit the diameter of its head was only four times the Earth's.

A comet's tail, lying always in the plane of its orbit, is even more enormous. The length of the tails of large comets is of the same order as the distance between the orbits of the inner planets. The length of the tail of Halley's comet of 1811 near perihelion (when its head is quite near the Sun) was greater than the distance from the Earth to the Sun. A light signal emitted at the nucleus of the comet would have been perceived at the end of the cometary tail only after ten minutes, while a shout, if it could be transmitted in the tail without weakening and at the same velocity as in air, would reach the tip of the tail only after twenty years.

The volumes of large comets are hundreds of times larger than the Sun's volume, and if the mean density

Comet Bennett photographed on 12/13 April 1970.



The distribution of comets with respect to aphelion distances suggesting the gravitational effect of the planets, notably Jupiter, on some "families of comets".

of comets were equal to the density of water, then the first invasion of the solar system by such a comet would produce utter havoc. All the planets would be torn out of their respective orbits, and the majestic Sun itself would begin to spin at breakneck speed around the comet or might even fall into it.

However, as they fly through the planetary system comets behave more like pale shadows silently flitting from planet to planet than terrible conquerors. The comet's smooth, easy pace, at which it sometimes even brushes against the planets, does not make the slightest impression on them, and on the contrary, every time a comet passes close to a massive planet it is moved away from its path. We have seen that this either transforms the comet into a short-period one or expels it from the solar system.

Owing to the absence of any perceptible attraction of the planets by comets it is not possible to determine a comet's mass and it is only possible to indicate its upper limit. If the mass of a comet were greater than that limit its gravitational pull on the planets would be noticeable, if only slightly. It so happens that the mass of even the very largest comet is less than 1/10,000 of the Earth's. Other evidence suggests that the mass of comets is even many times smaller – less than 1/1,000,000 part of the terrestrial mass.

From a cosmic point of view the insignificance of cometary masses is evident, but from the terrestrial point of view this mass is nonetheless immense. Even if a comet's mass were only 1/1,000,000,000,000 of the Earth's, it would still come to 1,000,000,000 tonnes, or about the same amount of earth as was excavated and hauled away during the digging of the Panama canal.

The small mass of comets, negligible as compared with the planets, with all the enormousness of their volume does not in the least hamper their movement in the solar system; not the slightest resistance is encountered there. Interplanetary space is airless, and if the Earth with everything on it were suddenly to disappear so that all that remained was one bit of down from the softest featherbed, this bit of down would continue to rush around the Sun at the same speed of 30 kilometres a second that it had moved when with the Earth. A comet's enormous "frontal section" in no way affects its motion.

It is apparent that the mean density of comets is fantastically small if you will just imagine the negligibly small mass of a comet spread throughout its enormous volume. Disperse 1/1,000,000,000,000 part of the Earth throughout a volume a hundred times larger than the Sun, whose volume in turn is 1,300,000 times the Earth's. Or, take a grain of wheat, separate out a millionth part of it and grind it into the finest dust, and then scatter this dust in the Bol'shoi Theater in Moscow – that will be about the mean density of comets. Can you imagine it? So then isn't it fair to call this distinctly visible comet "visible nothing"?

All our notions of how it is possible to "make a mountain out of a mole-hill" pale before the ability of nature.

Our calculations will be even more striking if we take into account that practically all the mass of a comet is concentrated in its tiny nucleus, that the enormous head accounts for incomparably less mass and the fantastic volume of the tail accounts for yet less mass. The densest part of a comet is its nucleus, followed by its head, then the beginning of the tail itself, which imperceptibly trails off into airless interplanetary space.

The author of the book first determined in 1943 the density of gases at different distances from a comet's nucleus (Comet 1943 I, with a mass of its gaseous envelope of $8 \cdot 10^4$ tonnes). The density varied inversely with the square of distance to the nucleus. Near it there were 10^{11} molecules per cubic centimetre, and at 370,000 kilometres there were only two molecules of cyanogen and one molecule of C_2 per cubic centimetre.

Only the nucleus of a comet can be hard. The entire mass of a comet's nucleus collected into a single solid body would be no larger in size than a small asteroid. If you think that comets, having heads and tails, ought also to have "trunks" as the main part of their mass, why then the "trunk" would most likely be the comet nucleus itself, located, moreover, inside the head!

The nucleus is the brightest and most opaque part of a comet. Faint stars are seen through the head, and even more so through the tail, and this indicates how tenuous the matter of a comet is. It is immediately apparent that the tail and head of a comet are formed from matter ejected from the nucleus, and the closer it is to the Sun the more intensive is the ejection. In periodic comets this matter is given off continuously, especially at perihelion, for hundreds of years and probably much longer. Consequently, during one revolution around the Sun only a negligible part of the matter in the nucleus is given off to form the coma and the tail.

The Cause of Luminescence of Comets and Their Chemical Composition

In Lomonosov's time nothing was known about the variation of comets' brightness, much less about their spectra. However Lomonosov, with characteristic scientific insight, described the luminescence of comets from a viewpoint very close to the modern one. He wrote, "The cause of comets' pale glow and of their tails is still inadequately known, which cause I am certain lies in the electrical force..."

If comets shone only by reflected light their brightness, as they approached the Sun (after allowing for changes in their distance from the Earth), would vary inversely with the square of their distance from the Sun. This is approximately the way the star-like nucleus of a comet behaves, which is in agreement with the view that it consists basically of hard chunks that simply reflect the Sun's light.

This is also confirmed by the character of the spectrum of the nucleus. Usually, it is a copy of the solar spectrum, as the spectrum of reflected light ought to be. But when a comet approaches the Sun bright sodium emission lines appear in the spectrum of its nucleus. In the spectrum of the nucleus of Comet 1882, which passed extremely close to the Sun, even the bright lines of iron and nickel were discovered, though they vanished as the comet receded from the Sun. Later the sodium lines disappeared. All this must be explained by the fact that when it approaches very close to the Sun the hard nucleus of a comet is heated to such an extent that it begins to vaporize, turning into an incandescent luminous vapour. Sodium vaporizes and gives off light at a lower temperature than iron does, i.e. at a greater distance from the Sun whilst closer to the Sun iron is affected, too. The distribution of brightness in the head of a comet resulting from these processes has recently been investigated by D. O. Mokhnach in Leningrad.

As it approaches the Sun the head of a comet varies in brightness much faster than inversely as the square of the distance, most often as the cube or even the fourth power of the distance. This indicates that the luminescence (brightness) of a comet's head depends on the Sun, but it is not simple reflection. Evidently the Sun excites the comet's luminescence, but it is a cold glow, i.e. it does not arise from the comet being turned into an incandescent vapour since a comet shines even when it is far from the Sun where its temperature must be way below zero Celcius. Dust cannot give this glow – only gases.

The brightness of a comet behaves in a rather fascinating way, and the above dependence on the distance to the Sun varies not only from comet to comet, but for one comet as it travels around the Sun. This undoubtedly implies that a cometary nucleus is unstable, and that its surface may change promptly. A striking example is the case of a comet discovered by the Czech astronomer Kohoutek in the early spring of 1973. At the time it was still very far away from the Sun and, therefore, very faint (the 16th magnitude). But its orbit, which was quickly calculated, appeared to have its perihelion extremely close to the Sun, only 0.14 AU or 21,000,000 kilometres. This inspired observers very much who, assuming that its brightness would be proportional to the forth or even a higher power of its distance from the Sun, expected that in December and January the comet would be nearly as bright as Venus and thus hoped to study it closely. The comet, however, grew in its brightness quite slowly and in December it was just visible by the naked eye, its observation being additionally hampered by sunrise. It was only in January of 1974 that it reached about the second magnitude and became susceptible

to investigation using average instruments. The fuss kicked up by journalists about the "comet of the century", as it was dubbed, then appeared a false alarm.

Certain molecules of cometary gas absorb sunlight and re-emit it later at the same wavelength. Physicists call this type of radiation *resonance* radiation. Other molecules absorb solar energy in the form of ultraviolet rays, but emit them in the form of rays of another visible wavelength. Physicists call this *fluorescence*. An example of fluorescence can be seen in certain substances on Earth, for instance zinc sulfide, which when "illuminated" by invisible X-rays in the dark gives off visible light, often green or blue. A theory of the origin of comet spectra was developed by Swings in Belgium and supported by the latest observations.

The spectrum of the comet head shows that it consists of molecules, i. e. chemical compounds, that emit broad *bands* rather than narrow bright lines. Only in recent years has it been possible to elucidate the chemical composition of these gases. It was found that the head of a comet consists of molecules of carbon (C_2), cyanogen (CN), and hydrocarbon (CH). Nitrogen hydride (NH), hydroxyl (OH) and NH_2 have been discovered recently.

Using an artificial satellite, a comet was first observed in 1970. It was found in ultraviolent light (that is mostly absorbed by the atmosphere) that the nucleus of Comet Tago-Sato-Kosaki 1969g was surrounded by a hydrogen cloud that is wider than the Sun. The enormous size of the cloud in itself was no surprise for astronomers because thirty years before the present author proved that the comet 1943g has an envelope of cyanogen vapour larger than the Sun.

The brightness of the various bands in the spectrum is different in different comets, and even in the same comet it varies with its distance from the Sun, apparently owing to changes in the proportion of the gases that make up the head of a comet, and also to changes in the conditions of its emission. The main role is nonetheless played by carbon and cyanogen, which is known to be extremely poisonous gas and the main constituent of a deadly poison – hydrocyanic acid.

The spectrum of a comet, in addition to bright bands, also has a continuous component, which also seems to belong to the gas molecules and is not the spectrum of reflected sunlight. The majority of scientists nevertheless believe that there must be dust in the head of a comet and that the bent tails of Bredikhin's type II must be made of it since they are also observed as having a continuous spectrum. If we succeeded in discovering the dark lines in this spectrum that are present in the solar spectrum, this would be compelling evidence for the presence of dust in cometary tails.

For the most part, however, the spectrum of a comet tail is gaseous, indicating the presence of carbon monoxide (CO) and nitrogen molecules (N_2). As you may know, carbon monoxide is formed in stoves and engines when the fuel is burned incompletely. It is quite poisonous too, but not in the same way as cyanogen.

You can see that questions about a comet's chemical composition cannot be answered briefly any more than can, for example, questions about the content of a large circus performance; the composition of comets is diverse and complex, and in the various parts of comets (nucleus, head, and tail) it is different.

What Occurs Within Comets

The eminent Russian astronomer Fyodor Bredikhin devoted the better part of his life to the study of cometary phenomena and to the formulation of a theory of comet tail formation. By the close of the last century he had originated a consistent theory now accepted by all scientists.

Luminous material is often seen gushing forth from the comet nucleus within the head of a large comet on the side facing the Sun. Sometimes it has the form of several jets. Headed at first in the direction of the Sun the jets turn to the sides, flow back, and skirting the nucleus create around the comet's head an envelope having parabolic outline. The gases of the head, discharged in great quantity, recede farther and farther from the Sun and create the comet's tail. The material of the tail at the same time constantly moves away from the Sun and, scattered into space, is replaced by new material from the nucleus. The closer the comet is to the Sun and the more strongly the nucleus is heated, the more rapid and massive is the gas discharge from the nucleus and the longer, brighter and more luxurient the tail.

The nucleus of tailless comets gives off too little material, which is also what large comets do when they are far from the Sun and have no tails. These comets are nearly round.

Occasionally, something like an explosion is

observed in the nucleus of a bright comet, because suddenly a bright cloud issues from it, quickly passing into the tail and moving along it at an accelerated pace. This cloud is sometimes elongated and lies across the tail. It happens now and then that a series of clouds is ejected from the nucleus one after another. Some of them move so rapidly that in a period of a few days, and sometimes even in a matter of hours, they run the entire length of the tail – tens of millions of kilometres – to be dispersed.

Bredikhin drew attention to the shapes of cometary tails. Some are nearly straight, pointed almost directly away from the Sun, and others are curved. He accounted for the curvature of the tail by the magnitude of the repulsive force coming from the Sun and acting on the particles of the tail. The greater the repulsive force compared with the gravitational pull of the Sun, the straighter the tail. Bredikhin derived formulas to calculate this force from the shape of the tail, and Orlov improved them and gave more accurate techniques of computing the forces from the movements of clouds in comets' tails.

In very strongly bent tails (type III) attraction was found to be greater than repulsion, in tails that are less curved (type II) repulsion and attraction are balanced and the ejected particles move by inertia, and in the almost straight tails (type I) repulsion exceeds attraction by a factor of tens and sometimes even of hundreds of times. The author of this book discovered particles in Comet Brooks 1893 IV moving under the action of a repulsive force 3,000 times greater than gravitational attraction!

A few comets have been observed to have bright cone-shaped appendages coming out of the head with the apex pointing toward the sun. The repulsive force does not act on them.

The relationship of the curvature of the tail to the ratio between the repulsive and attractive forces acting on it simultaneously may be visualized to some extent from the following illustration. Imagine a steam locomotive rushing along, emitting puffs of smoke from its stack the way a comet throws off particles from its nucleus. The warm air pulls the particles upwards, and the air resistance in calm weather pushes the column of smoke backwards. The air resistance increases with the speed of the locomotive. We liken it to the repulsive force of the Sun. When the locomotive is moving at high speed, the stream of smoke is deflected backwards straight away and does not curve, drifting into a straight line.

What is it that causes comet particles, which are most certainly being pulled in by the Sun, to be pushed away by it at the same time, and often with much greater force?

This was answered by another remarkable Russian physicist, Pyotr Lebedev (1866-1912). He was the first to prove by delicate experiments that light exerts a pressure on tiny particles. This follows from the theory of light. Light acts on any body with a force proportional to the surface of the body. For ordinary terrestrial objects this force is negligible by comparison with their weight and does not therefore have any perceptible effect. So the air resistance to a bomb dropped from an airplane is very small. But air resistance, which is proportional to the frontal area or cross section of a body, is very great compared with weight for a parachutist or a bit of down. They fall to earth much more slowly than the bomb because the area of the parachute and the down is great in relation to their weight or mass. Fine dust settles more slowly than coarser dust when a housewife is cleaning the room. The reason is the same – the weight of the dust particles is proportional to their mass, i. e. their volume or the cube of their diameter, while air resistance is proportional to the surface area of the dust particles, i. e. the square of their diameter. As the diameter decreases the weight falls off more rapidly than the surface does. If the diameter of a ball is reduced to one tenth of its original, its weight will be reduced to one thousandth of its original weight, but its surface will only be reduced to a hundredth of the original surface.

Therefore as the particles get smaller in size the light pressure on them will decrease more slowly than their mass. As the particles get smaller the ratio of the force with which light pushes on them to the gravitational pull of the Sun on them increases. As calculations show, when the dimensions of the dust particles are comparable to light wavelength, i. e. dimensions measured in thousandths of a millimetre or less, the repulsive force equals and may even become greater than the gravitational attraction of the Sun at any distance from it.

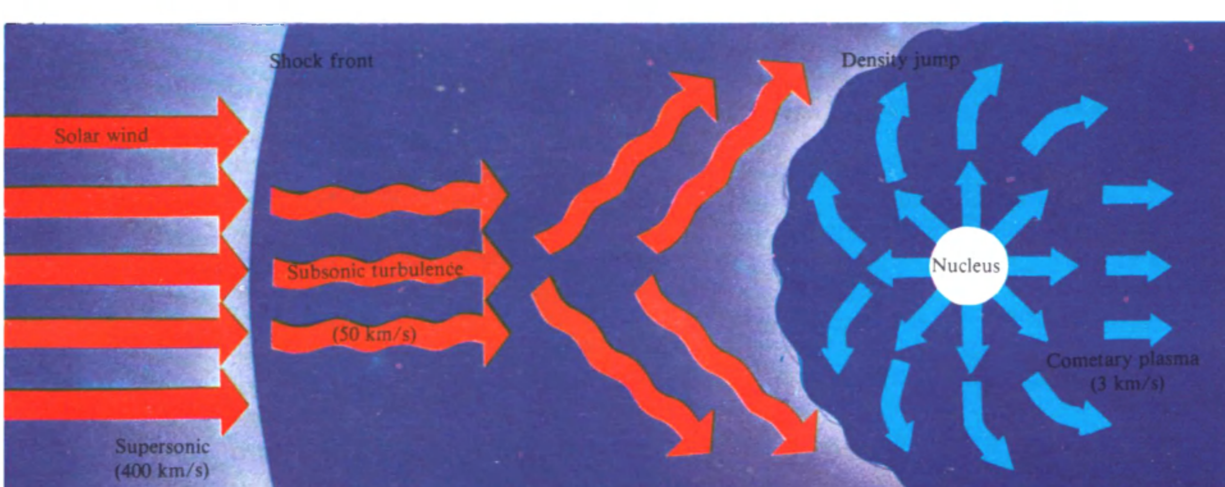
Thus comet tails of types II and III may be formed from extremely fine dust expelled from the nucleus, perhaps from collisions between the chunks of rock of which it may consist, or from dust that is liberated as

the "dirty ice" nucleus evaporates. But spectroscopic analysis has shown that in addition to dust, type I tails contain gases, chiefly ionized carbon monoxide CO^+.

Atomic and molecular theory tells us that light pressure acts on these particles as well, although molecules cannot be treated simply as small spheres. The force of light pressure on molecules depends on the molecular structure. Scientists have fairly recently succeeded in calculating this force for carbon monoxide molecules, and they found that light pressure on them is dozens of times stronger than solar attraction. Further development of this physical (not just mechanical) theory of comets awaits further progress in studies of the structure and properties of molecules. Many people are now coming to believe that electromagnetic forces and the bombardment of comets by particles ejected from the surface of the Sun play a part in the formation of cometary tails.

The Sun constantly emits streams of corpuscles, streams of *plasma*, i. e. jets of ionized gas carrying a magnetic field with them. The lines of force of the magnetic field behave like elastic threads. The ionized gases of the comet head, on colliding with the solar corpuscular stream, "press through" these lines of force and impart a catenary shape to them (this is the shape of a suspended heavy thread whose ends are fastened at the same level). The solar corpuscular streams and magnetic forces – the solar wind – may also play a part in the variations of type I tails, in the movements of cloud formation in them and in the other cometary processes that are still poorly understood. When the gas of the comet collides with the gas of the corpuscular stream, a shock wave is generated and the impulse of protons of the stream can be transmitted to the ions of the comet tail through the magnetic field carried by the stream. The tail gases cannot penetrate it and it draws them away. The molecules that comprise the tail are unfailingly ionized, but there are no such molecules at the head. At the same time the molecules of the head cannot give rise to the molecules of the tail through chemical reactions. It is therefore assumed that the nucleus liberates some unobserved "parent" molecules. Their bands may lie in the ultraviolet range unobserved from the Earth. When exposed to sunlight the "parent" molecules disintegrate and produce ionized molecules of CO and N_2 whose spectrum is visible to us. These molecules are affected by the solar wind. Abrupt enhancements of the "solar wind" must abruptly change the acceleration imparted by them and intensify the liberation of "parent" molecules. As to the causes of ionization of cometary molecules, they await further studies.

The filaments in type I tails, according to the Swedish scientist Hannes Alfvén may arise as the result of the propagation in them of a particular type of wave discovered by him in the laboratory. These waves are similar to the oscillations of an elastic thread and are the vibrations of the cometary plasma



The action of solar wind on comets.

together with the lines of force of the magnetic field "frozen" in them. This means that the gas can only move along the lines of force and the latter travel together with the gas.

Installation for comet modelling in the A. F. Ioffe Physico-Technical Institute of the USSR Academy of Sciences.

A Collision Between the Earth and a Comet

A collision between the Earth and a comet is what people began to fear once they had ceased viewing comets as harbingers of war. But to speak of a collision between the Earth and a comet would be about the same as saying that an unguided aerostat that had broken loose in London might accidentally fall in New York. It is extremely doubtful that the aerostat would be carried by the wind all the way to New York, and it is even more doubtful that the balloon would fall in the centre of a city. Here the probability of the aerostat falling in a field, a forest or a city is proportional to the area occupied by the fields, forests and cities on Earth.

As regards the Earth colliding with the hard nucleus of a comet, a nucleus that has approached the Sun to the same distance as the Earth has one chance in 400,000,000 of colliding with the Earth.

Since about five comets come this close to the Sun every year the nucleus of a comet may collide with the Earth once every 80,000,000 years on average. Such is the probability of a collision! It is equal to the probability that out of 80 million white balls among which there is only one black one, and taking one ball each year, you will pull out the black ball in any given year.

Collision with the head or tail of a comet may, of course, occur more often, even much more often. But what would happen then? Many fascinating novels have been written on this theme.

Some people imagine that a collision between the Earth and the tail of a comet would be something like a crocodile brushing an egg with its tail. In the light of what has been said about comet tails, there is no need to be apprehensive. A comet's tail could neither change the Earth's orbit not mar its surface. But couldn't we be poisoned by the toxic gases present in the comet – cyanogen and carbon monoxide?

Knowing the negligibly small mass of comets, almost imperceptible under simulated laboratory conditions, we are convinced that the concentration of comet gases in our atmosphere would be quite imperceptible. The latest methods of chemical analysis would probably be unable to detect it. Considering the high velocities of celestial bodies, the Earth would not remain within the head or tail of a comet for more than a few hours. The comet gases, being of negligible density, would only mix with the topmost layers of the Earth's atmosphere. Literally only a few molecules would be able, after a long period, perhaps years, to reach the lower atmosphere. Then it is not certain that they would survive such a trip since they would undergo many collisions and chemical reactions with air molecules.

As far as we can tell from computations the Earth once crossed the tail of Comet 1861 II and Halley's comet passed between the Earth and the Sun 24 million kilometres from the Earth on 19 May 1910. Its tail then extended 30 million kilometres, and apparently touched the Earth on 19 May. Not only did nothing special happen at that time, but even the most sensitive chemical analysis failed, as in 1861, to detect the presence of any extraneous gases in the atmosphere.

Thus a "collision" between the Earth and the tail of

a comet containing carbon monoxide is much less dangerous for the Earth as a whole than is closing the damper on a stove that still has unburnt coal in it. It is even a pity that even a rare encounter with a comet does not permit us to study the chemistry of comets directly! Hopefully an automatic station sent directly through a comet's tail will be helpful.

But what would happen if the Earth nonetheless collided with the nucleus of a comet? After all, it is hard!

As we know, the mass of a cometary nucleus is negligibly small compared with the Earth. Studies by the author showed 40 years ago that, to our good fortune, the solid matter in the nucleus is divided up into a great number of pieces, so that even the largest of them will probably be no larger than a small cottage. If we accept that the nucleus consists of a mixture of ice and dust, then during its flight through the atmosphere the ice would immediately vaporize, while the dust particles would do even less harm than would be done if the nucleus consists of small pieces of rock.

Most of the pieces comprising the comet nucleus must be even smaller otherwise there would not be enough surface area to release the gases at the observed rate. For big game, a charge of small shot is less dangerous than one large-calibre bullet. So for the Earth, too, the shot-like structure of comet nuclei would be preferable in an encounter with them. In addition, atmospheric resistance would slow the motion of small hard particles more than it would slow larger ones, thus weakening their striking force. These pieces, already some distance apart in space, would be dispersed even more while falling to Earth and would fall tens or even hundreds of kilometres apart, and not in a heap.

What would happen as a result? At worst there would be slight local earthquakes and destruction over some areas a few kilometres in size.

In the chapter "Heavenly Rocks and Dust" we will discuss at length the fall in Siberia in 1908 of a giant body that exploded in the atmosphere so that essentially no remains of it have been found. Academician Fesenkov maintained that this was a collision of the Earth with the nucleus of a small comet. If it had really been the case, the collision could only have been within the range of several kilometres. On the other hand, as we shall see in the next chapter, the nucleus of a comet eventually disintegrates to form a cloud of tiny particles. The cloud is thousands of times larger than the Earth, the particles being widely spaced. Should the Earth encounter such a cloud, it is most likely that after an explosion microscopic particles would fall out on the ground. Atomic, hydrogen and other bombs are a good deal more dangerous for humanity.

The probability of fragments falling on a city is very small. To convince yourself of this try walking in one direction by the compass for a thousand kilometres and count how many cities you would pass through on your trip, then work out how much of the trip you spent walking on the pavement.

Where Were Comets Born? Are They Born Now?

The issue of the origins of comets is a very complicated one. So far the evidence available is insufficient to solve it. But astronomers, like the general public, are impatient to learn this, no matter how poorly the subject has been investigated so far. Therefore, they construct a variety of hypotheses about the origins of comets changing and developing them as new evidence becomes available. This explains the differences in opinions.

Take, for example, the discovery of asteroids with elongated orbits resembling those of some of the periodic comets. These small asteroids differ in their appearance from such comets only in that they have no nebulous envelope around them. Such are the asteroids Hidalgo, Hermes, Adonis, Apollo, and Icarus. On the other hand some short-period comets, such as Comet Schwassmann-Wachmann and Comet Oterma, have almost circular orbits. Besides, in some of the comets the nebulous envelopes are barely visible. Up until fairly recently, therefore, it was thought that there was a relationship between comets and asteroids, and that comets might well be formed as a result of disintegration of asteroids in collisions, with the fragments continuing on their path along altered orbits. Some astronomers even went as far as to support that asteroids, such as Hidalgo, are cometary nuclei that had lost their gaseous envelope. But the evidence gleaned so far concerning the nature and structure of nuclei does not warrant the assumption that these are stone or monolithic debris, thereby complicating this hypothesis.

Many years ago Vsekhsvyatsky, a Soviet

astronomer, found compelling indications of periodic comets quickly decreasing in their brightness, with the result that they become depleted and go unobserved. We also have witnessed examples of comets disintegrating and turning into meteor showers. Meanwhile new periodic comets continue to be discovered. If during the lifetime of the solar system their numbers have not dwindled, this strongly suggests that they are made up due to the birth of new comets. But where and how?

Vsekhsvyatsky contends that comets keep emerging even now, being produced by volcanic eruptions on planets or their moons.

In order to escape from giant planets, and all the more to break through their atmospheres, these ejections must have enormous speeds and unthinkable energies. At the same time the mass of all the comets put together must be about that of the planets, let alone their moons. Therefore, the hypothesis did not have notable following. There are also doubts that the moons can have volcanoes that are still active. Recently, however, in the spring of 1979, Vsekhsvyatsky's hypothesis was given big boost. The American space probe *Voyager 1* flew past Io, one of the four major Jovian moons, and took pictures of phenomena occurring on its surface. They have been interpreted as powerful, multistream lava eruptions from five volcanoes accompanied by ejections of ash and gases up to 500 kilometres. The ejections were silhouetted against the background of the sky at the planet's edge. This is good evidence in favour of the comets' origin hypothesis and Vsekhsvyatsky's other hypothesis about the formation of rings around giant planets having large moons. True, the mass of possible volcanic eruptions from the moons of the known planets is insufficient to produce all the comets known in the solar system.

But cannot it be (and this is my personal view) that here, as in other cases known to science, different events yield similar results. For example, craters are produced both by meteoric impacts and by volcanic eruptions. It may well be that the short-period comets could come from eruptions on the moons, and the long-period ones from the processes at the edges of the solar system.

The indefatigable Vsekhsvyatsky has established in the USSR at Kiev the first comet observatory in the world. He has also organized regular observations in the USSR of all the comets coming into view and has been sending expeditions of his co-workers to all parts of the country. It is hoped that these efforts will contribute significantly to our understanding of the nature and origin of comets.

The Dutch astronomer Jan Oort thought that comets originate far beyond Pluto's orbit. It may be, he argued, that at first these were just small condensations produced during the formation of the planets and then thrown away by planetary disturbances to the far reaches of the solar system. And now, owing to the action of the nearest stars, some of them are pushed back now and then and become visible.

It is not clear in these concepts of the cometary cloud at the boundaries of the solar system why such small bodies are so rich in gases.

We are now witnessing a resurgence of interest in comets, whose "viciousness" has been disproved by science. New evidence keeps coming in from land-based and space observations, providing further clues to our understanding of cometary phenomena.

4. Shooting Stars and Meteoric Storms

Shooting Stars and Stones from the Skies

When a star shoots through the sky, leaving a light trail, some superstitious people say, "someone has just died". These people think that every man has his own star in the sky, that a fortunate man has a bright star, and that the possessor of a dim star has a dull life. This is why they think each time a star passes by a life has passed.

If this were so then in the days of great battles, when thousands of people were shorn of life on the battle field, the sky should be strewn with falling stars like a tree shedding its yellowed leaves on a windy autumn day. And if stars actually fell from the sky, there would now not be a single star left in it because a well trained eye can count no more than 3,000 stars in the entire hemisphere of the sky and even without a war on Earth far more than 3,000 people die every year.

Many various guesses and suppositions have been made about the nature of shooting stars, but for more than one hundred years it has been firmly established that falling stars, otherwise called meteors, are no more than small stones the size of a seed or less that hurtle from interplanetary space into our atmosphere and are turned into scorching vapours.

The skeptical reader would ask: "How could be such a complicated statement be proved?"

In the 4th century B.C. the Greek philosopher and scientist Aristotle considered meteors to be terrestrial vapours, and for the next two thousand years no one knew how to prove the assumption, which was in any case invalid.

It was not until 1798 that two German students Brandes and Benzenberg perceived that the nature of shooting stars would become much clearer, if the distance to them could be successfully determined, and they thought of a method to do this using the idea of parallactic displacement we have already discussed earlier in the book.

If a shooting star (meteor) passes relatively close to an observer, its path against the background of a starry sky varies depending on where the observer is located. Two observers separated by 30-40 kilometres will see a meteor as being at different points in the sky. The difference in the *apparent* path of a meteor will for them be the larger, the farther the meteor is away from them.

The observers had only to be sure that they had seen and drawn on a star chart the path of the same meteor, noting the time of its appearance, its brightness and colour. If they both saw, for example, a greenish meteor, as bright as a second-magnitude star, at the same instant of time, then it was the same meteor, and the difference in its apparent path in the sky, judged from its sketch on the star chart, would be the parallactic displacement.

From these observations it is possible to determine, trigonometrically, the distance to any point on the meteor's path (for example, the points where it disappeared). We will not bore the reader with the trigonometry or describe the details of a graphical solution of the problem.

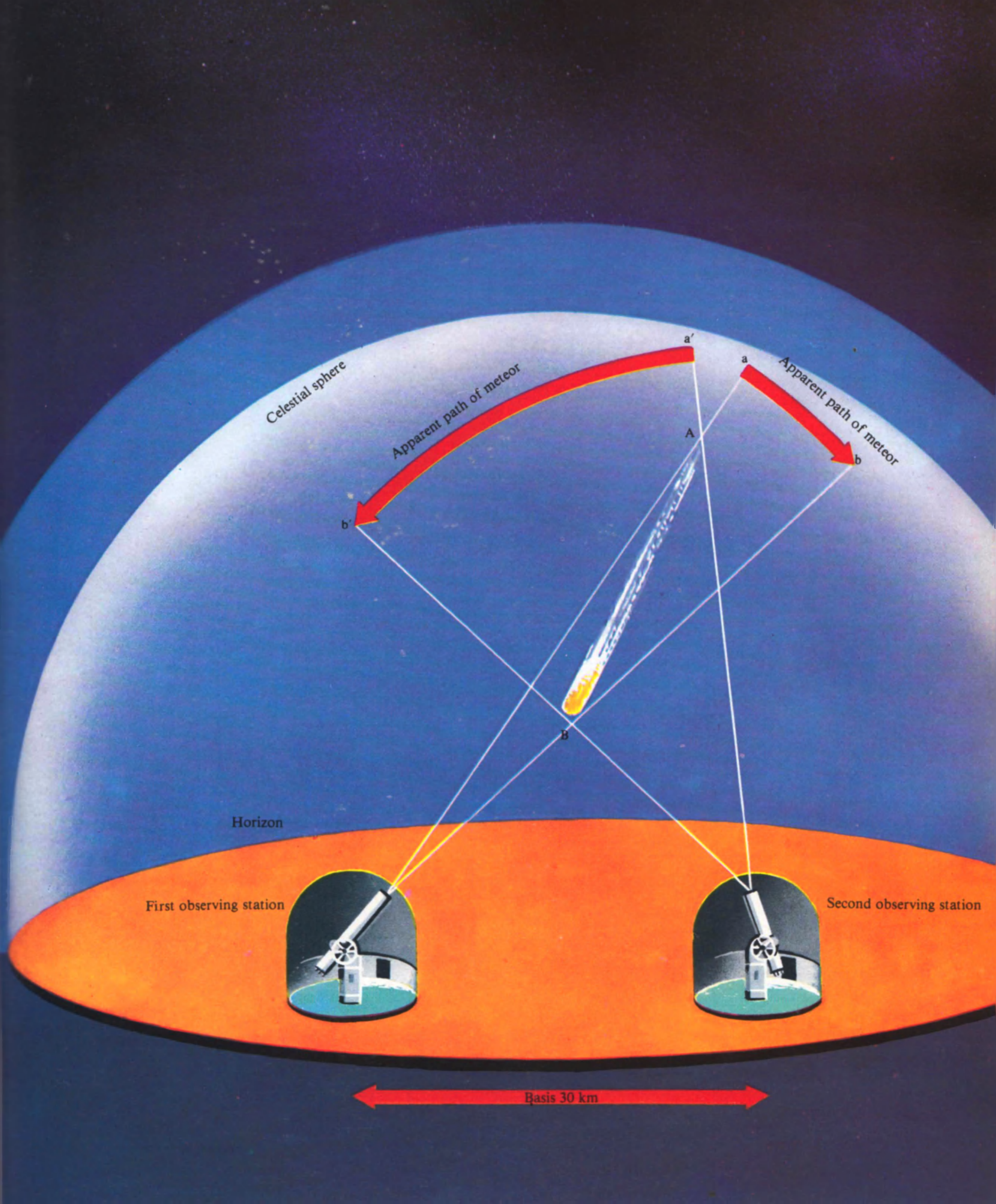
Although the meteoric events take place in the atmosphere (their heights above the Earth tell us this), they do not originate on the Earth (their speeds tell us this). They cover their entire path of 30-40 kilometres in approximately a second or less. Such a speed – tens of kilometres a second – can only be possessed by a body born in space and plunging into our atmosphere.

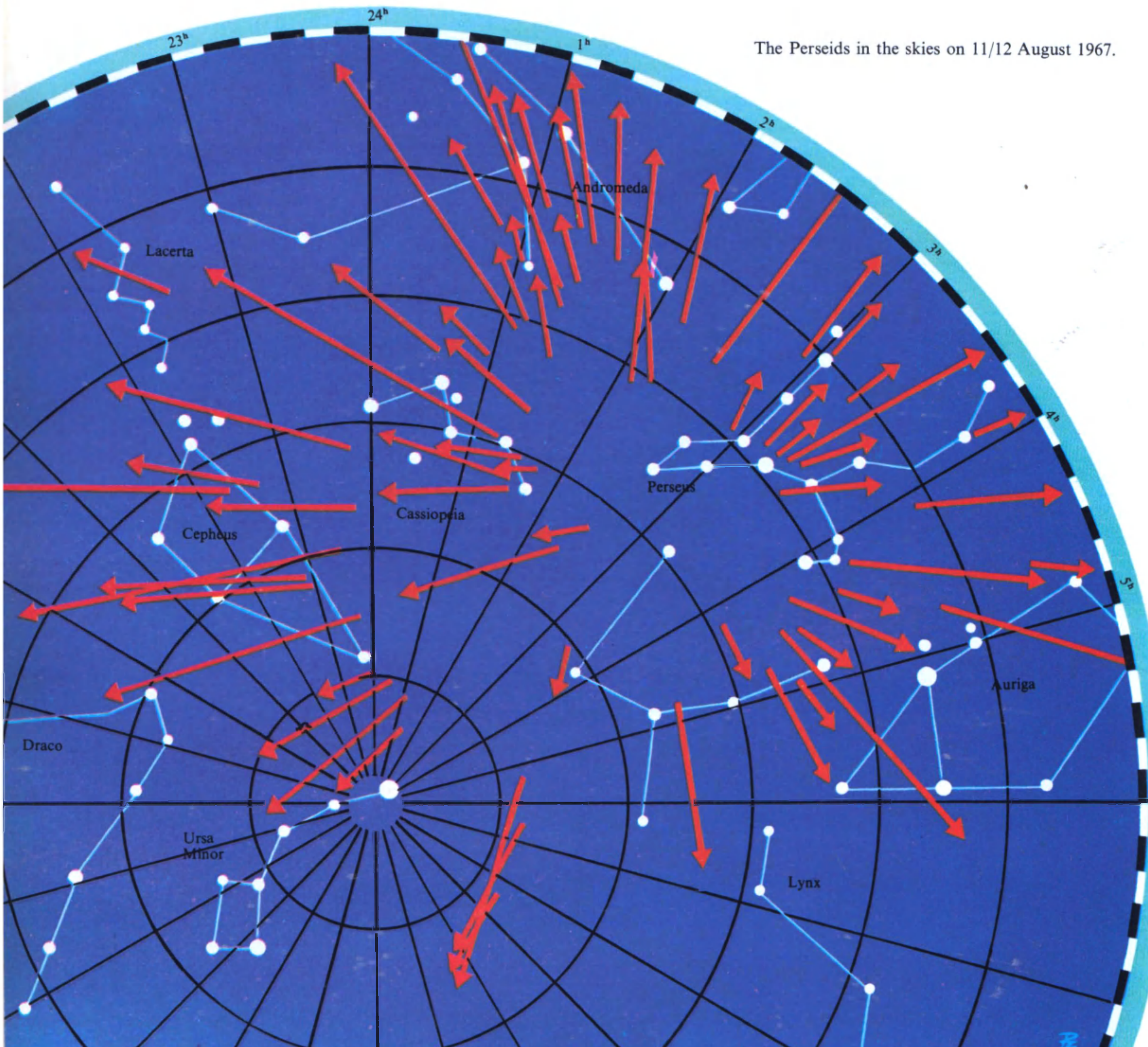
It was noted in 1833 that during a meteor storm, when meteors fell in profusion, all of them flew from the same constellation – the constellation Leo – in the evening, at night, and at dawn.

Consequently, the point of the sky from where meteors came, took part in the diurnal rotation of the sky, an apparent motion brought about by the diurnal rotation of the Earth. Thus the Earth turns relative to a stream of meteors, just as it does relative to the stars, and so meteors approach us from interplanetary space, they are not related to the Earth and do not take part in its revolution.

Meteors glow in the Earth's atmosphere, but they do not originate in it, and fall into it from outside, from space. Meteors are guests to our planets, strangers from airless space, but they are our guests for a shorter time than moths live on a summer day. Whizzing at tens of kilometres per second through the atmosphere, they become incandescent due to atmospheric resistance, turn into vapour and flare up for a few seconds, dispersing in the air. Such is the fate of these small stones, which have perhaps journeyed safely for millions of years in interplanetary space and

The determination of meteor paths.


Celestial sphere
Apparent path of meteor
a′
b′
a
b
Apparent path of meteor
A
B
Horizon
First observing station
Second observing station
Basis 30 km



The Perseids in the skies on 11/12 August 1967.

expired like moths flying into a candle at their first contact with the air covering the Earth.

The author has repeatedly let out the secret that meteors are small stones flying into the Earth's atmosphere. It is now the right place to state that this conclusion was first of all based on the cases of "stones from heaven" falling to the ground.

Among the meteors there are some extraordinarily bright ones – not just shooting stars, bur rather fast flying balls, sometimes visible even in day-time. They

are called *bolides*. Where a bolide ends its brilliant flight, a dark cloud appears in that part of the atmosphere, and a stone falls to Earth. Only the larger stones reach the ground, which have not been vaporized completely during their flight through the atmosphere. Stones falling to the ground from the sky are termed *meteorites*, they vary in size from a speck of dust to rocks the size of a large cupboard. Two American astronomers, Whipple and Jacchia, found not long ago that typically the solid fragments that produce meteoric trails have loose construction. That meteors are stones has been supported by recent photographs of meteor spectra.

Meteor Portraits and Passports

The first scientist at the beginning of the 20th century to photograph and study meteor spectra was Sergei Blazhko at Moscow University. Then still a young scientist, he directed a small wide-aperture camera toward the sky with a prism mounted in front of the objective lens.

Even simple photography of a meteor is a matter of rare success, because a meteor, as a quickly moving lighted point, weakly lighted at that, has no time to leave its own "portrait" on film. With modern sensitive photographic films only the very brightest meteors, brighter than second-magnitude stars, leave traces in the form of a thin line on the film. Flashes that may occur during the flight of a meteor appear in photography as a widening of this line.

It is necessary that a bright meteor flies past the region of the sky covered by a given camera. True, it is possible to select a night for this purpose, when there are many meteors; nevertheless it is still rare for a satisfactorily bright meteor to fly through the field of vision of a camera.

When astronomers wishing to photograph a meteor asked Blazhko: "What do you think are the chances of success?", he replied unfailingly with a joke; "It depends on good fortune and possibly on whether you are lucky."

The element of luck here is, of course, great, but the chance for success may be increased by using several cameras at once, pointing them at several regions of the sky. The several cameras will probably capture the celestial fugitive. Such a battery of cameras, watching for the appearance of meteors almost every night, is installed, for example, at the Dushanbe Observatory.

On average one meteor is obtained for every 100 hours of exposure (of course, not on one, but on many films) on a moonless night with an ordinary (e. g. small) number of shooting stars. In August, when many falling stars fly from the constellation Perseus, one meteor is photographed on average per 5 hours of exposure.

Would you have the patience to photograph meteors? Many amateur astronomers from the Moscow Branch of the All-Union Astronomical-Geodetic Society have the patience, and have been rewarded with a series of beautiful and valuable photographs. By the way, the equipment of the "meteor hunter" can be very modest – one or more ordinary cameras and a supply of film.

If a prism of the sort used in spectral photography is placed in front of the camera, instead of the image of a

A meteor trail on 13 October 1953, 23.07 EMT, over Jena (the radiant star Alpha U Mi.).

bright meteor its spectrum will be obtained. The spectral width is determined by the visible length of the meteor track along the sky.

Photographing meteor spectra is more difficult or requires more luck and time than ordinary photography, because the light of a meteor, stretched into a spectrum and absorbed in a prism, is much weaker. In 1908 and 1914 Blazhko obtained the spectra of several meteors and investigated them in detail. Only in recent years has it been possible to obtain new spectra of meteors by the joint efforts of astronomers from all over the world. At present only about fifty of such spectrograms have been successfully taken.

What was the result? The meteor spectra have shown – as their "passports" would – that some of them consist of rock, while others consist of iron with some nickel. But meteorites are exactly the same. Some of them are rocks, which only a specialist could distinguish from ordinary terrestrial rock. Others consist of pure iron with some nickel but in proportions never encountered in natural iron. From atrificially smelted iron they differ in unusually high crystallization, which we will discuss in more detail later. Now and then meteorites fall in the form of an iron sponge, the pores of which are filled with rock mass. These are stony-iron meteorites.

Meteor Storms and Streams

Small particles – we will call them pebbles, although many of them are iron – are derivatives of meteors, or shooting stars, and are frequently carried in large swarms. But the largest bodies move singly in space and fall to Earth singly, in any case not in company of meteors. It's as though they didn't want to know the small fry, and meteor storms and whole meteor showers observed are never accompanied by meteorites. This indicates that although the composition of meteors and meteorites is the same, their occurrence may be different.

Meteor storms and showers have been seen more than once, although rarely. The first of these to be described scientifically, was observed in November 1799. On 12 November 1833 a similar meteor storm was observed at night throughout the Earth. Stars fell as snow flakes in a winter snow storm, and frequently a spectator noted up to 20 meteors in one second. They fell so fast over the entire sky that one observer could not see them all. It was calculated that one on-looker, more or less vigilant, can follow an area of the sky with a diameter of 60 degrees. Because meteors glow at a height of about 80 kilometres above the Earth, the area of the atmosphere surveyed by our observer at this height was nearly 5,000 square kilometres. But the surface of the atmosphere is 100,000 times larger. On an average moonless night, when there is no meteor storm, an observer notices nearly ten meteors an hour. Accordingly in 24 hours it should be possible to count 24,000,000 meteors shooting through the terrestrial atmosphere.

Instead of 10, nearly 70,000 meteors per hour were counted on the November night in 1833. This means that hundreds of milliards of meteors flashed through the skies that night! And how many of them flew past the Earth! No one knows this exactly, because tiny pebbles that whizzed past the Earth were completely invisible, they did not glow. It is only possible to say that the number of pebbles flown past, probably, was as many times larger than the number getting into the Earth's atmosphere and portrayed by a fiery tail, as the number of rain drops in a strong downpour is more than the number that falls on the head of some unfortunate man in the storm.

The last especially mighty meteor storms took place on 9 October 1933, when up to 300 meteors per minute came from the constellation Draco. The meteors were not bright and, appearing in the night sky of Europe, toward midnight they had almost ceased, so that, when night fell in America there was no trace of this beautiful show. A swarm of pebbles flying to Earth was carried past it faster than the Earth had time to make a quarter of a turn about its own axis.

It is possible, on certain days of the year, to see an intense meteor fall, although it cannot be called a meteor storm. For example, every year on 9-14 August a large number of meteors fly from the constellation Perseus, on 19-22 April from the constellation Lyra, on 9-12 December from the constellation Gemini, and so on.

The centre of a place from which meteors fly in all directions, like arrows, is called the *radiant*. Meteors, whose radiant is located in the constellation Perseus, are called Perseids; those whose radiant is in Lyra Lyrids, and so forth.

It is not difficult to understand that the convergence of the meteor trails to one place – the radiant – when

traced backwards is perspective. As a matter of fact, the meteors at the end of their path approach us, which is why they seem to diverge. Their visible paths start higher up in the atmosphere farther away from us, and at a great distance their paths seem to be close together. You have seen many times how railroad tracks merge in the distance due to perspectives, while actually they are parallel to each other all the way. Parallel flying particles in a shower begin to glow as they crash into the air and as they approach the observer they seem to diverge ever more and more from the point (radiant), from which they are carried to us.

Every day the Earth encounters separate meteoric particles and every night individual meteors are visible that cannot be attributed to any shower. These are called sporadic meteors. In comparison with shower meteors they are like soldiers who have strayed from the ranks.

The distance between the particles of a meteor swarm can be worked out from the number of shooting stars observed in an hour in a given area of the atmosphere and the speed of the particles. It appears that meteors fly in space in loose swarms. For example, in a Perseid shower a volume of 10,000,000 cubic kilometres is apparent for each particle. So the average separation between particles is no less than 200 kilometres. If we were to encounter them beyond the confines of the Earth, for example flying in an imaginary interplanetary ship, we would not suspect what magic celestial fireworks they might produce for an observer on Earth.

More About Meteors

Meteors and meteorites are fascinating from several points of view and they should have a little more attention devoted to them.

First, meteorites are singular celestial bodies that fall into our hands. Of all the things in space only meteorites can be studied directly. You may touch, measure, crush, and analyze them, i. e. treat them as another terrestrial object. All the other celestial bodies must be investigated indirectly: by observing their apparent positions and motions and analyzing their light. To a nonspecialist the results of these studies often seem doubtful, approximate and far-fetched. That is why he is not completely satisfied, although actually many of these calculations are far truer than, say, our assumptions about some areas of our own planet, like for example arctic regions or the rain forests of Central Africa. In any case it is not true to say "as inaccessible as the stars in the sky".

Another reason why meteors and meteorites attract attention is that they are closely related to other celestial bodies and phenomena (comets, asteroids, zodiacal light, the solar corona, and the so-called dark nebulae in interstellar space) and they also cause certain types of surface relief to form on several celestial bodies, including our Earth.

These pebbles are partly the debris of some objects destroyed in a catastrophe, partly the "building blocks" from which various bodies, and maybe even our Earth, were formed.

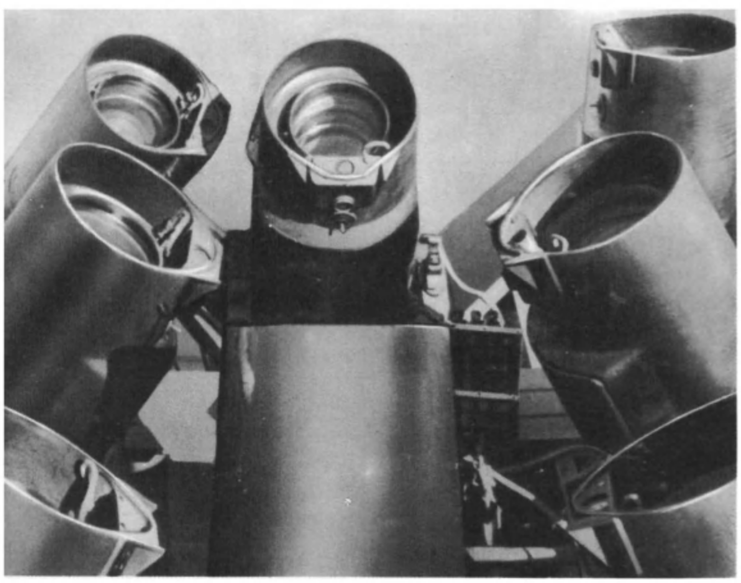
Finally, the evidence from meteors and meteorites can be used to study the upper layers of the Earth's atmosphere, in which scientists, aircraft designers, radio operators and even artillerymen are interested, but which until recently were inaccessible for continuous study. We will turn to this later, but now we will treat meteors as objects from space, if only the smallest of them.

What is it about meteors that interests us, and what remains to be found from observation? Well, we are interested in the heights at which meteors appear and disappear above the Earth's surface, their velocities, the dependence of velocity magnitude on meteor brightness, the number of meteors at various hours of the day and time of year, the distribution of their brightness and size, their path in space before encountering the Earth...

One of the greatest Soviet "hunters" for falling stars, I. S. Astapovich, recorded nearly 40,000 meteors for 15 years of work.

Anyone can observe meteors for science because the majority of meteor observations are done with the unaided eye and do not require any specialized knowledge. And even instruments for meteor observation in the majority of cases can be so simple and modest that any amateur can afford them.

Amateur astronomers such as Denning in England have played notable roles in meteor science. In the USSR an entire organization of amateur astronomers under the All-Union Astronomical-Geodetic Society is occupied with meteor observations. This organization plays a big role in the development of our knowl-



Equipment to photograph meteors at the Astrophysics Institute of the Tajik SSR Academy of Sciences.

edge about meteors and has in its possession an extensive archive of observational data. Such organizations function in many countries. They have also come to study meteors now in observatories, notably in Czechoslovakia, the USA, the USSR, and, by radio methods, in Great Britain.

We have already talked about how to determine the height of various points of a meteor trail (the method of intersection). Two observers, separated by 30-40 kilometres, simultaneously scan one and the same region of the sky and sketch the paths of meteors among the stars on a star chart. Comparing their sketches and identifying meteors common to both observers according to the time of observation, brightness, colour and approximate location, they can measure the perspective (parallactic) path shift of the meteor, that is how one observer saw it in comparison to how the other saw it.

The end of a meteor's path, which is closest to the observer, is displaced more than the beginning. Knowing the height of the start and finish of the meteor's path in the atmosphere and the projection of the path on the Earth's surface, it is not difficult to establish its true length. By estimating the duration of the flight of the meteor in the Earth's atmosphere and dividing its path length by it, you obtain its *average* speed since the actual speed of the meteor in the atmosphere is not constant: it varies due to the braking action of air resistance.

Above all the speed of a meteor in the atmosphere is of interest to us because knowledge of it helps us to find where the meteors come from – from interplanetary or interstellar space.

We know that the speed of the Earth about its almost circular orbit around the Sun is 30 kilometres per second. The theory of gravity states that a body moving along, as far from the Sun as the Earth is and travelling at a speed not exceeding $30\sqrt{2}$ kilometres per second relative to the Sun, i. e. not exceeding 42 kilometres per second, cannot overcome the Sun's gravity. It then moves in an ellipse, periodically returning towards the Sun, and is thus a member of the solar system.

At speeds slightly greater than 42 kilometres per second, a solid body's original path is only bent by the Sun's gravity, which does not close it. The body thus travels round the Sun in a hyperbola and pulls away from the region of its attraction. When it approaches the Sun in a hyperbola, it is the first and last time it appears in our solar system, and so obviously comes from interstellar space where the attraction of our Sun is weaker than that of other stars.

Parabolic motion occurs when the body moves at 42 kilometres per second and so borders between elliptical and hyperbolic motions. It is practically impossible. Should such a speed arise accidentally, the attraction of the planets would immediately increase or decrease it, so changing a parabolic orbit into an elliptical or hyperbolical orbit.

You shouldn't imagine that 42 kilometres per second is an absolutely fatal number. According to the theory of gravity, for every distance from the Sun there is a velocity V_0 at which the motion of a body must be circular; this motion will be elliptical at a velocity greater than V_0 but less than $V_0\sqrt{2}$, and hyperbolic at a velocity of only slightly more than $V_0\sqrt{2}$. When a body is travelling at less than V_0 for its distance from the Sun, the body will fall into it along a curved line. A body will fall onto the Sun in a straight line, if its speed is equal to zero, i. e. if the body is stationary relative to the Sun, it falls freely, like a seed into a field.

We only observe meteors near our Earth, and there-

fore the number 42 kilometres per second is actually "fatal", as it were, for the observed meteors. Unfortunately, simple sketches of a meteor's path do not give the desired precision and result in speed overestimates, due to the difficulty in remembering precisely the path, particularly its beginning, because of the suddenness of its appearance and because of the transiency of appearance.

Greater accuracy is provided by using two cameras separated by several kilometres from each other. In that case, however, a section rapidly turned by an electric motor must be placed in front of the camera lenses, so that the camera objective lens is covered several times a second, and exposure is interrupted several times.

As a result, a meteor's trail is interrupted on the film, the interruptions having equal durations though unequal lengths. It this way it is clear that the meteor moves faster at the start of its path and slower towards the end. This brings out the braking effect of the atmosphere. Meteor height is determined in the same way as with naked-eye observations – from the parallax of the meteor path against the background of the stars, fixed on both photographs.

Unfortunately, these paired pictures are even rarer than ordinary ones.

Meteor brightnesses, in naked-eye observations, are compared using the standard adopted for stars, and so astronomers talk about meteors of first, second, etc., star magnitude.

Observations indicate that the brighter the meteor, the deeper into the atmosphere it penetrates, but that the height of its flash point is almost independent of its brightness. An overwhelming majority of meteors begin to glow at an altitude of 100-120 kilometres and extinguish at 80-85 kilometres. It turns out that at this height there is a peculiar layer in the atmosphere where the air density increases abruptly. This layer – an invisible air barrier – destroys the remainder of the meteor. So the majority of our heavenly guests perish at this "wall", crashing into it.

It can easily be seen that for a given velocity, which determines the force of air resistance, and hence the rate of its vaporizing and brightness, a meteor will be the brighter the larger its mass. Only the more massive and slower meteors penetrate the "armour" at nearly 80 kilometres and pass through to below, where they break up entirely at a height of 30-40 kilometres. Bolides reach this height where they produce a sound that is often remembered as a hissing. Finally, meteorites falling to Earth ordinarily stop glowing at a height of nearly 22 kilometres and thence fall to Earth as dark, unlit bodies with the usual speed of free fall. At this point their reserves of cosmic velocity are usually exhausted.

On the other hand, the greater the speed of meteors at their penetration into the atmosphere, the greater is the height at which they begin to glow and disintegrate. At high speeds the resistance of the air increases in proportion to the square of the speed, and possibly faster. Therefore, a meteor with a speed of 20 kilometres per second glows at a height of nearly 60 kilometres, and with a speed of 70 kilometres per second it starts glowing at nearly 100 kilometres.

A knowledge of a meteor's speed relative to the Earth is important for atmospheric and meteor studies. But to study the origin of meteors, it is necessary to know their speed relative to the Sun. Their speed relative to the Earth is their speed relative to the Sun plus the Earth's speed. For example, a meteor flying directly to the Earth at 40 kilometres per second, plunges into our atmosphere at a speed of 70 kilometres per second, because the Earth itself adds 30 kilometres per second to its speed. Such a meteor in pursuit of the Earth rushes up to it with a velocity of only $40 - 30 = 10$ kilometres per second, though the attraction of the Earth will slightly increase this velocity.

Since at any moment of time the magnitude and direction of the Earth's velocity is known, it is possible to subtract it from the observed velocity of a meteor to obtain its velocity relative to the Sun. It is necessary here to take into account the angles between velocities and path deviation and the meteor's speed as it is affected by the attraction of the Earth.

Pictures taken using a revolving slotted disk in front of the camera definitely indicate that the velocities of meteors are clearly elliptical, i. e. that meteors are permanent members of the solar system. Three such meteors seem to have orbited around the Sun (until they were destroyed in the Earth's atmosphere). Their periods are on average about 4 years and their orbits have a major semiaxis of 2.5 astronomical units and an eccentricity of the order as that of the periodical comets and several asteroids (0.7), and lie nearly in the plane of ecliptic.

A velocity of 42 kilometres per second at one astronomical unit away from us is the velocity of a body alien to us, moving in a parabola. Whereas a speed of 41 kilometres per second (only 2-3 per cent smaller than the critical one) already corresponds to a period of 27 years in an orbit that is nine times larger than that of the Earth.

Thus the tiniest error made in determining a meteor's speed (and it is extremely hard to determine!) leads to a new conclusion about the meteor's place in the solar system.

Recently the velocities of thousands of meteors outside the Earth have been determined using a completely new method. The reflection of the radio waves from the tracks left behind by meteors has been observed. A great many meteor observations have been made by radio techniques at Jodrell Bank, England, and at Dushanbe, USSR. These very precise observations showed that practically all meteors move in elliptical orbits and are members of the solar system. Perhaps a few solitary meteors approach us from outside, like rare guests. Thus the mystery about the origins of meteors has been solved conclusively.

The theory of meteor glow has yielded the following meteor masses. The mass of a very bright meteor of zero star magnitude, given its velocity in the atmosphere is 55 kilometres per second, is 0.25 gramme. This is equal to the weight of several drops of water. The mass of a meteor of fifth magnitude, barely visible to the eye, is several thousandths of a gramme.

Since we can work out the masses of meteors, their dimensions too are no longer mystery. An ordinary bright meteor before its disintegration in the atmosphere is the size of a pea, and weak meteors, visible only in a telescope, are the size of a small pin-head (these masses and sizes are approximate). Such bodies are so far from being like real stars, yet uninformed people only differentiate them by the epithet "shooting".

It may be that doubt may arise as to how such crumbs are visible to us being hundreds of kilometres away? But then the shooting star visible to us is not this hard particle! It is an unusually bright incandescent vapour, into which it is converted in the atmosphere. The vapour forms an incandescent gaseous atmosphere of a significant size around a flying particle. It should also be remembered that the very fine incandescent filament of an electric lamp, thanks to its glow, is visible from far away, even though it is hundredths of a millimetre thick, besides, the gases into which a meteor is converted are heated much more strongly.

It is no wonder, therefore, that a bright meteor will be visible hundreds of kilometres away like a star of second magnitude and will have an actual light power of 3,360 standard candles.

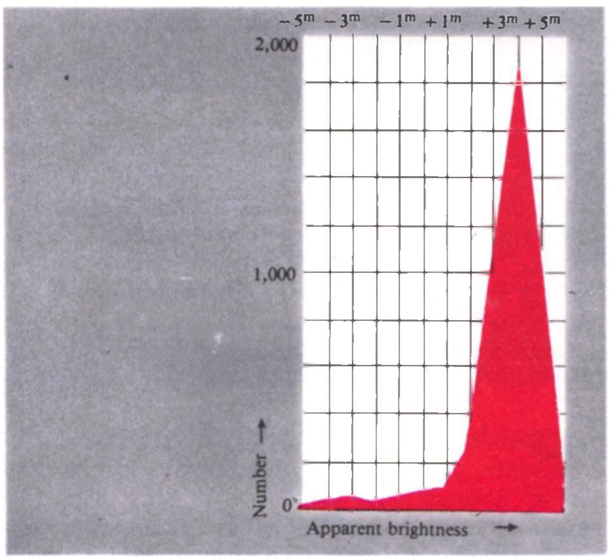
Very small cosmic specks precipitating onto the Earth are the miserable remains of fairly large pebbles, the majority of which vaporize during their flight.

A Census of Meteors

If there are people who consider it impossible to count the naked-eye stars, they would consider it even more hopeless to try and count every shooting star that becomes visible anywhere over the Earth during a single year. They have in fact been counted, though not one by one of course. Suppose we want to know the number of trees in a particular sector of forest, it does not matter very much to us if we skip a hundred or so trees in the count, we will be satisfied knowing that there are, say, nearly 10,000 rather than nearly 3,000 or 170,000. Moreover, our curiosity will be particularly satisfied by this approximate number, if before this we didn't have the slightest idea of the forest's size. For example, it is curious, although hardly important, to know that an average man, not recognizably bald, has nearly 200,000 hairs on his head. Before we know this we would only be able to guess how many there are, i. e. several thousands or millions. Our idea will change very little, if we were a thousand or so hairs out in our count, or our estimate was in error even several times worse.

Thus, by counting the number of meteors of various apparent brightness at various hours of the same day, and repeating this several times a year, it is possible to estimate how many of them fall in a year. Knowledge of this number not only satisfies simple curiosity, it tells us much more. In particular, it may indicate by how much the Earth's mass increases due to meteors and what percentage the meteoric matter accounts for in the composition of soil. What if suddenly it appears that the potatoes in our kitchen garden grow in a layer formed by centuries of accumulation of meteoric material.

When counting meteors, the percentage of meteors



Distribution of meteors in apparent brightness.

Number and Mass of Meteors, Daily Hitting the Earth

Apparent stellar magnitude	Number of meteors	Meteor mass, milligrammes	Total mass of meteors of given magnitude, tonnes
−3	28,000	4,600	0.11
−2	71,000	1,600	0.11
−1	180,000	630	0.11
0	450,000	250	0.11
1	1,100,000	100	0.11
2	2,800,000	40	0.11
3	7,100,000	16	0.11
4	18,000,000	6,3	0.11
5	45,000,000	2.5	0.11
6	110,000,000	1.0	0.11
7	280,000,000	0.40	0.11
8	710,000,000	0.16	0.11
9	1,800,000,000	0,063	0.11
10	4.500,000,000	0.025	0.11

not noticed by an observer must be taken into account by comparing the simultaneous observations of several people, the area of the atmosphere they scan, and the meteors only visible through the telescope.

The results of such a count are presented in the following table, from which, by the way, it is obvious that with a weakening of the stellar magnitude of meteors, their number increases 2.5 times per unit. However, a decrease by one magnitude implies a decrease in brightness by 2.5 times, and its mass decreases accordingly (because for a given speed meteor brightness is proportional to its mass). Thanks to this fortuitous coincidence, the total mass of meteors of every stellar magnitude is the same, namely 110 kilogrammes.

As we shall see, the "efficiency" of meteors, if we view them as light sources, is very great. If all meteors belonging only to a single stellar magnitude category falling in one day fell simultaneously through your field of vision, they would generate several times more light than the full moon, and if all meteors falling in a day flashed all at once, they would light the ground 250 times stronger than the moon. And all of this is achieved by converting only 5 tonnes of material into incandescent vapour hundreds of kilometres above the Earth's surface! If they glowed one kilometre above us, the light would be 10,000 times brighter, if only for one second.

The very brightest of meteors, or bolides, have a brightness corresponding to the −10th stellar magnitude. On the other hand, the number of weak meteors, not visible even with a telescope, should not be taken to be infinite.

Meteor mass decreases with brightness, and meteors weaker than the 30th star magnitude are so small that such dust particles would have been long ago blown away from the solar system by light pressure alone, which for them exceeds the gravitational pull of the system.

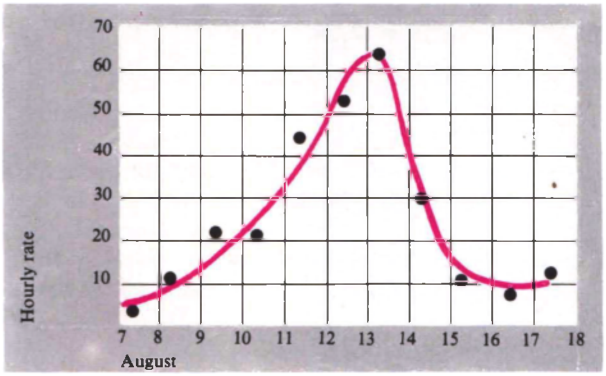
Thus the total mass of meteors from −10 to +30 stellar magnitude, falling to Earth daily, is 4,400 kilogrammes. Similar data for meteorites yield 5,500 kilogrammes more. The daily total of meteorite material falling to Earth is nearly 10 tonnes.

Assuming that since the Earth's crust hardened, i. e. approximately two milliard years ago, meteors and meteorites have fallen as frequently as they do now, then ten thousand tonnes of meteorite material would have fallen on every square kilometre of the Earth's surface, a layer less than 10 centimetres thick. Therefore, although meteorite material is mixed into the soil, there is an insignificant amount of it, and there is no good reason to say that our gardens grow on meteoric soil.

Meteoric Swarms

Up until now we have chiefly been discussing sporadic meteors. We will be concerned now with meteor showers, i.e. meteors falling on fixed days of the year

Shower	Date of maximum number of meteors	Velocity relative to Earth, km/s	Discription of meteors
Quadrantids (from Boötes)	January 3	46	Yellowish, long trails
Lyrids	April 21	51	Fast, white with trails
Eta-Aquarids	May 4	66	Ditto
Delta-Aquarids	July 28	50	Yellow
Perseids	August 12	61	Fast, white
Orionids	October 22	68	Ditto
Leonids	November 16	72	Fast, greenish
Geminids	December 12	36	White



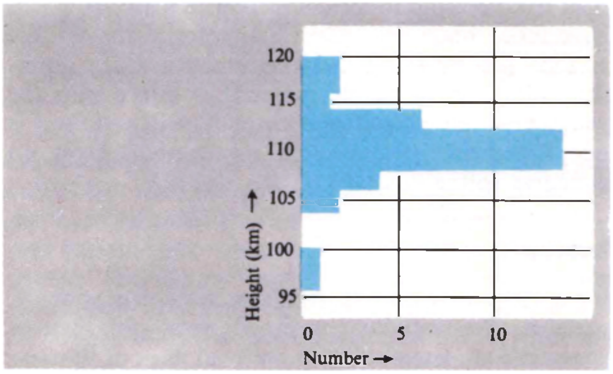
The average hourly rate of the 1967 Perseids at about maximum intensity.

and flying from a specific radiant. The table above is a list of the showers richest in meteors.

Even the fact that meteors, coming from interplanetary space and flying from a specific radiant, are early observed on the same days, suggests that they move in some sort of orbit in a shower. On these days the Earth crosses their path and brushes through them. If the meteors are more or less stretched out in an orbit, one behind the other at regular intervals, the Earth will cross their paths to collide with them on approximately the same date each year. This is the case with Perseids. The Earth crashes through the swarm of Perseids for several days. On 12 August, obviously, it intersects the middle of this stream, where the meteors are the thickest and the shower is at its maximum.

It is easy to see that if meteors move in an elliptical orbit around the Sun, and are unevenly distributed along their orbit so that they bunch up together in some places, the Earth will intersect the parts of the orbit poor in meteorites more frequently, and in these years (always on the same days) they will be more scarce. Then the Earth meets the main crowd of meteors on their yearly path, there will be an abundant rain of stars.

Height distribution of the Perseids.



If a meteor swarm has a very short duration, then it will not meet the Earth every time it arrives at a point where the two orbits intersect, because either the swarm or the Earth will have passed the intersection point earlier than the other, as if they were playing hide-and-seek. It is quite obvious that such cases do occur, but we do not know of them as yet. Actually, the orbital periods of most showers must be measured in decades, and the Earth and shower will meet once in several centuries. By the way, meteor studies are only about one hundred years old.

Given the precise coordinates of the radiant and the velocity of the meteors, it is possible to work out the orbit a meteor shower has in space. This orbit varies with time as the attraction of the planets, notably Jupiter, perturbs swarms into other orbits.

It is impossible to imagine meteors distributed over an orbit cluster with time. It would be expected, on the contrary, that the attractions of the planets and the Sun, exerting different actions on nearby and remote sections of the orbit will gradually disperse the swarm in all directions, mainly along its orbit, so that with time the swarm will be smeared out over the entire orbit.

It is clear that the longer a meteor swarm revolves about the Sun the longer it will be subjected to "pullig-apart" action and the wider it will be stretched out in

orbit. So it should be possible to judge the age of the swarm by the concentration of meteors in their orbit, its period and orbital orientation relative to the planets playing a role, of course.

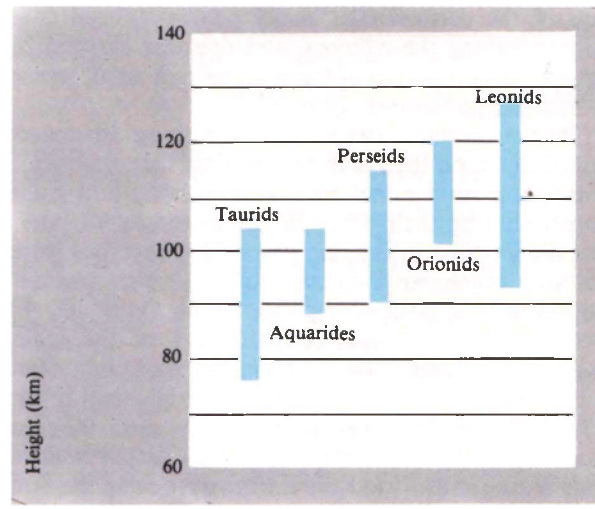
It is possible that sporadic meteors are just "lonely wolves", particles extracted somewhere from a large company of objects similar to them.

Many meteor showers have not only ancient origins, but also ancient record of their appearances. Leonids are the most remarkable in this respect. They have an orbital period of 33 years, and whole showers of meteors from this radiant were observed, for example, in 1799, 1833, and 1866, the Earth meeting their clusters every 33 years.

A meteoric storm produced by the Leonids was seen and described for the first time in South America in November 1799. The Indians were shocked by this event and remembered it full well, whereas the Europeans seem to have overlooked it completely.

It is on this basis that the periodic nature of meteoric storms was first suspected and, in fact, the November storm of shooting stars recurred in 1833. Scientists turned to the chronicles of various peoples and have traced, although with some gaps, the Leonids as far back as to 1768 B.C. The first record was made 3,700 years ago by Chinese chroniclers. The next occurrence was found in Arabic sources dating to the year 902. Japanese chroniclers noted unusual star falls in the Novembers of 867, 1002, 1035-1037, which even frightened Japanese emperors into declaring amnesties to prisoners. Later, chroniclers in various lands more and more often, not without superstitious fear, noted maxima of Leonid showers. So the ancient Russian evidence in the famous Laurentian chronicles is among those of interest. The notes of 1202 say: "At 5 o'clock the whole sky fell, the stars fall rapidly from the sky, all of the stars will fall to earth". In 1533 it says that in Moscow "many people saw stars pulled from the sky, flying from the east to the wintry west." In another annal this is described as a wonderful "apparition in the heavens", as a "vision" by Sexton Tarasy from the bell-tower in Novgorod the Great: "a multitude of angels shooting fire arrows like rain strong from a cloud".

According to the American scientist Fisher, the Russian chronicles as a whole contain evidence concerning meteors for several centuries that is valuable for science and is not to be found in Western Euro-



Average heights of major meteor showers.

pean records. Thus, ancient observations have come to be of use for modern science.

The Leonids forecast for November 1866 were observed everywhere, but in 1899 they failed to appear in quantity. It seems that between 1866 and 1899 the meteor swarm passed near Jupiter and Saturn. The pull of these planets seems to have distorted the orbit so that the Earth only met the edge of the swarm. Again, in 1932 the Leonids did not live up to expectations and made a meagre appearance of only one meteor per minute, as in 1899. It is unlikely that a planetary disturbance will at some time direct this shower to out planet, as it is just a tiny speck in space where there is so much room for meteoric orbits.

Leonids approach the Earth almost head on, colliding with its "morning" hemisphere, and their velocity, added to the orbital velocity of the Earth, results in their speed of 72 kilometres per second in the atmosphere. At such a high speed they vaporize in the air very rapidly and brightly, leaving trails in the form of quickly decaying misty arrows.

If, however, disturbances from Jupiter and Saturn have almost deprived us of the spectacular show of Leonids, every now and then these very disturbances present us with new surprises. From the unknown younder are coming fresh wonders. The most recent of these were the Draconids which are referred to in "Meteor Storms and Streams". In spite of the sudden-

ness of the performance, many people succeeded in photographing the meteors, and one film showed 26 meteors over a period of ten minutes in a small sector of the sky measuring $10 \times 10°$.

The Draconids, having rambled long in space, rained down on the Earth for the first time that year. Until then Jupiter had been continuously and stubbornly shifting their orbit, finally causing it to intersect the Earth's. In ensuing years the Draconids were not seen much, an indication of their high concentration in one place along their orbit. On 9 and 10 October 1946 we again touched part of the main swarm and again saw a storm of shooting stars.

In 1933, natives of the Sudan were frightened by an evil spirit "tearing stars from the sky" and tried to scare it away by drumbeat, just as the Chinese tried to scare a dragon away to keep it from devouring the Sun during an eclipse.

In astronomy only one dragon is preserved, and that is a constellation; we rarely remember the reason why it was called so long, long ago when superstitions reigned.

Meteoric storms have scared the populace more than once. So in 1833, Negro slaves on American plantations took the Leonids to be an omen for doomsday, and one hundred years later, in 1933, the Draconids terrified people in the then backward Portugal, inducing them to flock to the churches.

Cometary Dust

While they may terrify ignorant people, meteoric storms may by contrast help astronomers to dispel absolutely different fears, those, so to speak, of a scientific nature, namely the fears that the Earth may collide with a comet. Comet Biela and the meteors of the same name, the Bielids, are especially valuable for allaying these fears and also for clarifying the origins of meteors.

We have already mentioned that during its appearance in 1846 this comet broke in two, and the resultant, now much fainter, comets separated from each other significantly and returned to the Sun in 1852. Then they vanished into the thin air. Their subsequent approaches were calculated very precisely but they failed to appear.

On 27 November 1872, twenty years after the mysterious disappearance of Comet Biela, the sky sparkled with shooting stars, albeit not very bright ones. Their radiant lay in the constellation Andromeda. Their rate, like that of the Draconids of 1933, grew from late twilight to 8.30 p. m., when it reached its peak of hundreds of meteors a minute. But after midnight the sky only sparkled with a few rare meteors. The orbit of these meteors, as computed from the observed position of their radiant, was shown to deviate from that of the lost Comet Biela. The meteors seemed to fly in a row along the path the unhappy comet had moved before it disappeared. The link between them was so obvious that astronomers wondered whether the meteors were tiny but innumerable members of the Bielids (or Andromedids), the remnants of the tailed comet. We are led to believe this is so, although the comet had started degenerating before 1852. Closer investigation showed this actually to have been the case. As far back as 27 November 1782 an abundant meteor shower was observed that, obviously, was similar to the one just described. In 1832, the orbit of Comet Biela passed the Earth's orbit only several thousand kilometres away, although it did not meet the Earth. Up to 1872 the Earth and comet played hide-and-seek: either one or the other was late for the date. After 1852 the comet did not approach Jupiter closely, therefore it or what remained of it must have been circulating in the same orbit. An encounter between them and the Earth was possible, but their place in the orbit was not known exactly and so it was impossible to predict the time when they would meet.

On 27 November 1872 when meteors put on a show in the sky, the comet was already far away, many hundreds of thousands of kilometres, as it had crossed the Earth's orbit 80 days earlier, on 9 September.

In 1878 the Earth appeared at the point of intersection half a year earlier, and in 1879, half a year later than the comet. Few meteors were seen in those years. After another period expired, Comet Biela should have intersected the Earth's orbit in mid-January 1886, yet somewhat earlier, on 27 November 1885, a meteoric shower rained down. It was observed among others by the artist and amateur astronomer A. M. Vasnetsov (brother of the famous artist V. M. Vasnetsov). He related that the meteors appeared to be fairly lazy, with intervals of nearly half a second, leaving pale, dim trails, and they themselves were not bright, mostly about the third star magni-

tude. Their sluggishness is explained by the fact that Bielids often overtake the Earth and enter the atmosphere with a speed of only 20 kilometers a second.

We see that the swarm of 1885, stretched only a little at the time, preceded where the comet should have been located. The meteoric swarm that produced the storm of 1799 must have been originated earlier, perhaps in 1772 when the comet approached Jupiter, too.

Consequently, the comet's nucleus took no less than one hundred years to disintegrate into meteors (it was the nucleus because it alone contained solid matter), and the meteors began to form long before the comet as such was destroyed.

In 1890, and then again in 1901, Jupiter disturbed the motion of the Bielids and therefore very few meteors were seen near the time of their possible meetings with the Earth (Novembers 1892 and 1899). They have not been seen since. As their path now lies several million kilometres away from the Earth and, unseen by us, they tear along past and will never, it seems, come within a region accessible to our examination.

But the Bielids, a striking example of the connection between meteors and comets, were not the first step to this conclusion.

Back in 1866 the Italian astronomer Giovanni Schiaparelli (1835-1910) discovered that the orbit of Perseids is close to that of Comet 1862 III. Another weak comet was observed in 1866 and the resemblance between its orbit and that of Leonids was immediately noted by three astronomers in different countries. Before long, a special investigation uncovered a resemblance between the orbit of Lyrids and the orbit of Comet 1861 I.

So far nearly a dozen meteor streams have been tied to comets. So Halley's comet is related to Eta-Aquarids and Orionids.

Of great interest is the recently established connection between Taurids and the most short-period of comets, the famous Comet Encke. These meteors, whose period was found to be 3.3 years, seem to have separated from Comet Encke about 10,000 years back, and are now moving in an orbit that, owing to perturbations, differs markedly from the comet's orbit.

The table below lists the meteor showers that are known with certainty to be associated with comets, and some of the elements of their orbits. The table shows that the orbitary elements for "related" meteor streams and comets are close to each other. This listing, however, only contains a small portion of the meteor streams known so far. We will explain the reason for this.

Orbits of Meteor Showers and Parent Comets

Showers and comets	Period, years	Perihelion distance, AU	Eccentricity	Inclination, deg	Velocity relative to the Earth, km/s
Perseids	110	0.97	0.96	120	61
Comet 1862 III	122	0.963	0.96	124	
Leonids	33.2	0.986	0.905	163	72
Comet 1866 I	33.2	0.977	0.906	163	
Lyrids	–	0.90	1.00	80	51
Comet 1861 I	415	0.92	0.98	80	
Bielids	8.3	0.855	0.86	12.1	17
Comet Biela	6.62	0.861	0.86	12.6	
Eta-Aquarids	–	0.60	0.97	162	66
Orionids	–	0.52	1.00	161	68
Halley's comet	76	0.59	0.97	162	
Draconids	6.6	1.02	0.71	31	20
Comet 1933 III	6.6	1.00	0.72	13	
Taurids	3.3	0.39	0.82	4	27
Comet Encke	3.3	0.33	0.85	13	

Many of the observed meteor showers can be related, though with some uncertainty, either to comets that have long since disappeared, or to comets having a revolution period of the order of one hundred years or greater. It must be remembered that a comet may not be observed each time it approaches the Sun. Not infrequently its path is randomly located relative to the Earth, so that in most of its approaches to the Earth and to the Sun a comet stubbornly hides in the rays of the Sun. It also happens that the light of the full moon interferes with a comet's appearances, sometimes short-lived. It is possible that from the interplanetary deep suddenly arises a periodic comet unheard-of before, and its relationship with some well known shower becomes established very quickly.

It must be confessed that in the cited examples of the establishment of connections between meteor showers and comets astronomers were lucky. As soon as the orbits of three meteor showers had been first determined, three weak comets appeared and there was a lucky coincidence. The comet 1861 I, the ancestor of the Lyrids, has a long period, and its next appearance will be in several hundred years. Comet 1862 III, which scatters Perseids in its orbit, also next time returns to the Sun at the end of the current century. The fellow traveller of the Leonids, Comet 1866 I, although revolving about the Sun with period 33 years, was located in the sky so near the Sun in 1899

and 1932 that it, an old acquaintance, was not seen.

These three comets then passed unnoticed, as is often the case with weak comets, so that we might not have detected a relationship between meteor showers and the comets and might not have known that they existed in the same orbit.

It is possible to imagine astronomers puzzling over the occurrence of meteors and, perhaps, making fantastic hypotheses. We see that success or failure in science, besides systematic and arduous labour, at times also depends on fortune. We should not, however, forget that a luck alone is not enough, it is necessary to know how to make use of it and to be ready for it.

What has been said about the lucky occurrences of comets may also be applied to meteors. The data summarized in the table above have been gleaned for a century. In this century the Bielids supplied not a single meteor, whilst the Draconids, one of the richest meteor showers, are a fairly recent development, and just a few years ago one line in the table was missing.

The years 1916 and 1921 saw the appearance of a few faint meteors associated with Comet Pons-Winnecke, but in 1922 and 1923 the author spent many a June night waiting for them in vain. The vigil went unrewarded and there were no meteors. Barely accessible to observation, they were again removed from the Earth's orbit by planetary perturbations, so that this addition to our list remains questionable.

In comparison with the number of comets ploughing through the solar system the number of comets noted is not large, and even more trivial is the number of the comets known to be accompanied by the products of their disintegration, meteors. Not a single comet moving in huge ellipses with aphelia lying beyond the confines of Pluto's orbit has shown us its meteors.

Rarely approaching the large planets, they seem to have decayed little, and their meteors are spread insignificantly along their orbits. Or if spread, they are so separated that, colliding with the Earth "in installments", in various years and so sparcely, they do not enable their radiant to be identified and hence their orbit to be worked out.

The perihelia of most of the comets, even the periodic ones, lie outside the Earth's orbit, and unless subjected to exceptionally intense perturbations, their meteors are not destined to meet with the Earth – ever. Few of the comets with orbits intersecting the terrestrial orbit are short-period ones and thus fall apart rapidly to yield observable meteors, but the very same disturbances that produced and scattered these meteors quickly take them out of the observation.

We have already said that particles in meteor swarms are separated by hundreds of kilomêtres. Knowing the Earth's diameter, the time the Earth takes to cross the meteor shower, and the number of meteors hitting the atmosphere, we can estimate the total mass of the meteors in the swarm.

The mass of the Perseids is estimated to be 500,000,000 tonnes. The result strikes us, but it is negligible compared with typical masses of heavenly bodies. To form the Earth 1,000,000,000,000,000 such meteor swarms would be required and one such swarm would only suffice to put down a 1-cm layer of dust over, say, Sicily.

Meteor spectra now available in fair quantities indicate that all the meteors arriving in periodic showers are stony, whereas the sporadic meteors are about equally divided between the stony and iron ones.

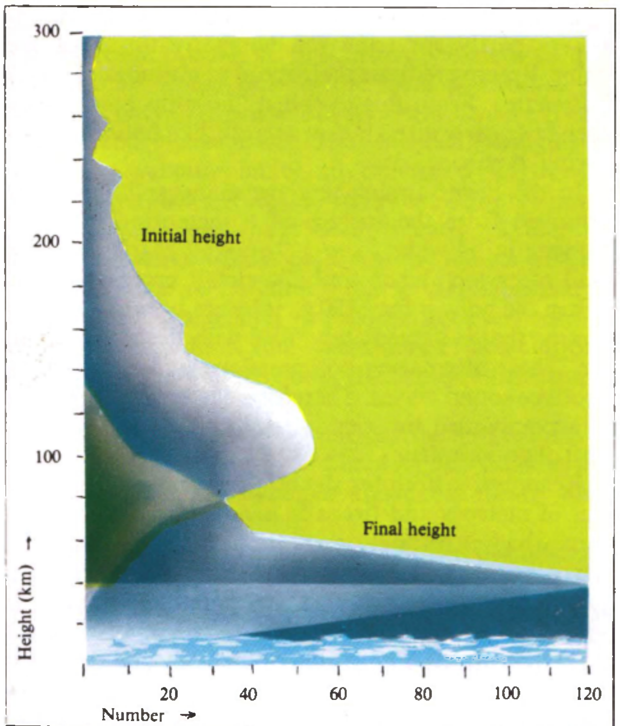
It might seem that this does not warrant the assumption that all the sporadic meteors like periodic meteor streams are derived from comets and are parts of dispersed meteor swarms. It may be that some of them, namely the iron ones, have other origins. It is impossible to estimate their total mass in the solar system. Anyway meteors and even meteor swarms travel in profusion throughout the solar system, and it is only a trifling fraction of them that is accessible for our observation.

Observing meteors we to some degree are like insect collectors walking along a narrow path across an enormous stretch of grassland and collecting only the samples that just happen to come our way.

But still we must be satisfied even by as little as establishing an undeniable connection between meteors and comets.

Meteors in the Atmosphere

We have thus debunked shooting stars as genuine stars, those great heavenly bodies, and have identified them only as insignificant stony or iron bodies. While beyond the terrestrial atmosphere, they are celestial, if trifling, bodies, and examination of them as such carried us into the depths of interplanetary space, and



Initial and final heights of bright meteors.

caused us to turn to that other, far more significant thing, comets. Once they hit the terrestrial atmosphere and glow in it for a short while both meteors and meteorites cease to be celestial bodies. Their flight through the air is accompanied by singularly fascinating display and the small piece of rock, or meteor, ceases to be such. This is why some astronomers suggest to refer to these stones as meteoric bodies, and reserve the name meteor for just the glow as it shoots through the atmosphere. It seems to me that this is not really necessary as it may lead to inconveniences of its own. We will now dwell a bit on why and how meteors become visible in our atmosphere and how their study contributes to our understanding of our home planet.

A star wheeling soundlessly through the sky, fragments of a far comet and shelling of peaceful towns in the rear of a battle, what could they have in common?

The year is 1918... German armies strive for Paris, but they are yet far away. Everybody knows there is no sign of the enemy closer than 100 kilometres from the beautiful city, certainly no grounds for panic. And suddenly... large shells begin to explode in the outskirts of Paris. What is the matter...? Where is the enemy?

It turned out that the Germans had produced super-long-range cannons firing at 120 kilometres. They shot 120-kilogramme shells from a 37-metre barrel with an initial speed of 1,700 metres per second at an angle of 55° about the horizon. This was the secret of the unusually long range. Quickly cutting through the lower denser layers of air the shell penetrated the thinner upper layers of the Earth's atmosphere, to reach far into the stratosphere 40 kilometres up. The thinner air there put up a much smaller resistance to the shell's motion, and instead of several dozen kilometres, it was thus hurled at a hundred. Admittedly, the German shelling was not especially accurate, they were counting on panic.

The inaccuracy of their shelling might be accounted for by the impossibility of predicting exactly the ballistics at such heights. At the time data on the density, composition and dynamics of the air at the height involved were scarce, the atmosphere of these layers had not yet been studied. As a matter of fact, even stratospheric balloons, later used to raise people and scientific instruments, had then only reached 22 kilometres, and unmanned balloons with recording instruments had reached 30 kilometres. Rockets reaching higher than 100 kilometres only began to be launched after World War II.

In earlier times any useful understanding of the processes at higher altitudes could only be gained by examining events observed there, and so meteors that pierce these upper layers every day have been one of the best indirect techniques. It is only fairly recently that the armamentarium of astronomers has been supplemented by such a powerful tool of many-sided investigation of the upper atmosphere as artificial satellites. That is why the extensive studies of meteors was an important item on the programme of the International Geophysical Year (1957-1958).

Meteors are unintentional explorers of the stratosphere and we have to be able to "debrief" them. The following is the results of this debriefing started just about 40 years ago.

Meteors plunge into the atmosphere at a speed a hundred times faster than the initial velocity of a gun bullet. Kinetic energy is known to be equal half the product of the square of velocity by mass. This energy

is wholly expended in heat and light radiation, in atomizing the body, in decomposing the molecules of the air and body into atoms, and in ionizing the resultant atoms.

The molecules and atoms of a solid body, including a meteor, are arranged in a certain order and form the so-called crystal lattice. As a meteor cuts into the air at tremendous speed, molecules of the air are forced into the lattice of the meteoric body. The farther into the terrestrial atmosphere the meteor gets, the denser the air it encounters and the more severe the bombardment air molecules subject it to.

The front of the meteor eventually gets a shower of such impacts that air molecules find their way into the meteor as a shell into a reinforced-concrete fortification. This "shelling" of the front surface disrupts the bonds between the atoms and molecules of the body, destroys the crystal lattice and rips away individual molecules of the meteoric matter which accumulates in disorder on its front face. Some of the molecules decompose into atoms, and some of the atoms lose their constituent electrons to become ionized, i. e. electrically charged. The electrons torn away from atoms pass the ions and are sometimes captured by "vacant sites". In the process, according to the laws of physics, the atoms emit light. Each species of atoms radiates its own wavelength, and so the meteor spectrum is a bright-line spectrum inherent in the glow of thin gases.

The deeper into the atmosphere, the quicker the disintegration of the meteor and the brighter its glow. At a height of less than 130 kilometres it becomes intense enough for us to see it.

Air molecules, too, suffer from the blows, but they are stronger than those of the meteor and are less frequently ionized. Besides, they are not so highly concentrated, therefore their contribution to the glow is so weak that the lines of the gases of which the atmosphere is made up (mostly oxygen and nitrogen) are not found in the meteor's spectrum.

Farther down in the atmosphere, before the frontal surface of the meteor forms a "cap" of compressed gases supplied partly by the gaseous products of meteor decomposition and partly by the atmospheric air compressed by the motion of the meteoric body. The streams of compressed hot gases skirt around the body tearing new particles away and speeding up the disintegration.

Larger meteors penetrate deeper into the atmosphere partly surviving the harsh treatment of the entry. Braking reduces their speed at altitudes of 20-25 kilometres. From this so-called "hold-up point" they then fall nearly directly downwards, like bombs from a diving bomber.

In the lower atmosphere the many solid particles removed from the surface of a meteoric body and flowing in its wake form a "smoky" black or white trail often seen when fireballs streak across the sky. When the body is fairly large it leaves behind it a rarefaction that is immediately filled with the surrounding air. These alternating compressions and rarefactions produce sound waves. Therefore, the entry of fireballs is accompanied by sounds resembling cannon shots and thunderbolts.

Responsible both for the brightness and for the colour of meteors and fireballs is not the hot solid surface, which is triflingly small, but the evaporated solid particles. Their colour therefore depends not so much on temperature as on which of the emission lines in its visible spectrum are the brightest. The latter depends on the chemical composition of the body and the emission conditions which are determined by its speed. As a rule of thumb, a reddish colour corresponds to lower velocities.

This is just a general picture of how modern science explains why meteoric bodies glow in the atmosphere.

Let us briefly sketch some aspects of these phenomena revealed quite recently in connection with stratospheric studies. For example, the examination of meteor braking casts some light on how air density varies with height. Clearly, the larger the air density, the stronger the drag, but the latter is also dependent on speed and body configuration, whereby airplanes, cars and even locomotives are generally streamlined. A streamlined body has no sharp corners and is so designed that as it moves fast air flows around it with the least hindrance possible.

Artillery shells experience high drag during flight. Meteors fly through the air ten times faster than shells, and so they encounter higher drag. From a picture of a meteor taken once in Moscow by amateur astronomers, members of the Astronomic-Geodetic Society, using a camera with a shutter rotating in front of the objective lens, a braking (sometimes called negative acceleration) of nearly 40 kilometres per second per second was deduced. This is 400 times the

acceleration of free fall under gravity. And this was 40 kilometres above the ground, where the air is so thin that a man would quickly die from suffocation.

Sound only occurs in air of a certain density. There is no sound in free space. Just as a bell is ringing in vain in a vacuum under an evacuated jar usually shown at physics lectures, so world catastrophes in the free interplanetary space occur silently. A colossal nova explosion or stellar collisions (by the way, highly unlikely) occur so noiselessly that if you happened to be near the catastrophe you would not even turn to look if it was behind you. The nature of noise in fireballs tells us very much about the density of the upper atmosphere.

Good opportunities for studying the air currents in the upper atmosphere are offered by the trails left in the sky at 20-80 kilometres above our heads after bright meteors and fireballs.

The "life" of a track depends on the illumination and the amount of matter converted into the fine suspension of dust. Air currents play a role as they drag dust particles aside, so "covering" the trail of a fireball. There have been cases when a fireball trail was seen for 5-6 hours.

The silvery trails seen at night after fast and bright meteors have another nature. They are gaseous and are always higher than 80 kilometres. Along the meteor path air molecules are ionized by molecular collisions at enormous speed, and by ultraviolet meteoric radiation. In the cylinder formed in the wake of the meteor ions slowly recombined with electrons – slowly because at the height involved the air is extremely thin and charge carriers are widely separated and so they have to cover long distances to meet the carriers of the opposite charge. The recombination is accompanied by radiation at appropriate recombination lines. At the same time ionized molecules move outwards with the result that the trail width increases. This, of course, makes the trails fainter, but some trails (normally seen for several seconds only) live sometimes even for an hour.

Meteors incessantly ionize air at heights from 80 to 350 kilometres, so maintaining there ionized layers. But the main cause of them is the ionization by light (ultraviolet) and corpuscular radiations (fluxes of charged particles).

Maybe, not all of you know that it is thanks to these layers that we can communicate in shortwave range. Radio messages beamed by a transmitter strike these layers at an angle and are reflected as if from a mirror due to the layers electric conductance. They do not stray into free space but, having been reflected downwards almost unattenuated, they are received somewhere else very far from the transmitting station.

The reflection of a radio wave is related to its wavelength. It is possible to probe the ion density in an electrically conducting atmospheric layer by changing wavelength and noting the wavelength at which the transmission is cut off, i. e. when the radio waves will not be reflected but will escape from the atmosphere. Radio sounding enables us to measure the height of the layers, which varies somewhat.

As might have been expected, it was found that variation of the meteor rates and even individual fireballs affect the transmission of short waves by causing short-term changes in the electric conductivity of the air at heights of 50-130 kilometres. Significant interferences in radio communication were noted, for instance, on 9 October 1933 during a Draconid shower.

So radio communication responds in an unexpected way to the appearance of the transitory remains of comets, bodies that seem to be of no consequence for our everyday life on Earth.

About a century ago a Moscow astronomer, V. K. Tserasky, accidentally noticed in the summer unusual luminescent clouds flowing in the northern part of the night sky. These could not have been ordinary clouds, which float no higher than 8 or at most 12 kilometres above the ground because the Sun could not have reached them from beyond the horizon with its rays, illuminating them so brightly. They must have been unusually high clouds. Actually, a comparison of sketches of their positions against the background of the stars drawn simultaneously from two different places (by Tserasky and Byelopol'sky) enabled the former to show for the first time that these clouds travel 80-85 kilometres up. Ever since they have been observed many times always in the summer and always in the northern part of the sky, near the horizon, for it is only at this height and only under these conditions that the sunlight can illuminate them from beyond the horizon.

These nocturnal luminescent (or noctilucent, as they are sometimes called) clouds are invariably at 82 kilometres. It may well be that these clouds, lying near the

lower boundary of meteor extinction, are formed by ice crystals frozen onto dust particles.

At 80 kilometres, where the air, it would seem, should be very "clean" (just remember the proverbial cleanness of mountain air!), there is some dust – well and good. But what would you say if somebody would tell you that there were a metal atmosphere above our heads!

We have rightly rejected the naive notions of old about "celestial firmament" above our heads and now we are about to suggest the idea of... a metal heaven.

Indeed, the spectroscope in the hands of the French astrophysicists Cabannes, Dufay and Gauzi showed in 1938 with devastating calmness that the night-sky spectrum always displays the well-known yellow line of sodium and the lines of calcium. In addition to these metals, they hope to find in the atmosphere aluminium and even iron. (Note in passing that to obtain a spectrum of a night sky, which seems to be black anyway, i. e. emitting almost no light, requires many hours of exposure.) The metals found in the atmosphere reside about 130 kilometres up and, of course, form no firmament. The trace quantities of the above metals are to be found among the molecules of the very thin air at these altitudes. Metal atoms seem to scatter in the atmosphere as meteors evaporate, and they glow colliding with other particles. Really, in some way or other, the products of evaporation of heavy elements must not only stay in the atmosphere, but also accumulate there. Whether they glow there or not is a question in its own right, but there are no reasons for them, having scattered at a height of about one hundred kilometres, to descend at once to the ground.

Thus, meteoric matter is everywhere. It lies under our feet, it continuously travels in space, and it hangs overhead.

Meteoric evidence has contributed to our understanding of the stratosphere. Not all of the inferences, for example the early conclusions of Lindeman and Dobson, are indisputable in the very young science of meteoritics, but they still serve as illustrations of the possibilities that open up. Based on their theory of luminescence of meteoric bodies concentrating on the air-body interaction, the above authors in 1923 explained the height distribution of meteor extinction points and concluded that at about 60 kilometres the air is strongly heated. They estimated the temperature there to be +30°C, and later even +110°C. It thus turned out that the temperature at that height appeared to be higher than the boiling point of water and at stratospheric air pressures water boils at less than 100°C.

This result came as somewhat of a surprise because the direct temperature measurements obtained up to 30 kilometres showed that it first dropped drastically with height, and at 11 kilometres (the lower boundary of the stratosphere) a layer began that had a nearly constant temperature of −50°C, irrespective of the season and climate. Strictly speaking, the stratosphere even behaves "the wrong way round": in winter, even in polar regions, its temperature is about −45°C, whereas in summer, even in tropical regions, it is about −90°C. A characteristic of the troposphere, or the lower atmosphere, is the temperature falloff with height, and at the equator it is higher (up to 15-16 kilometres) than at poles (9-10 kilometres). It is this upper boundary – where temperature stops varying with height – that marks the beginning of the stratosphere, accounting to a certain degree for the unexpected distribution of stratospheric temperatures over climatic zones, since the stratosphere temperature equals that of the upper boundary of the troposphere. But seasonal and unexpected variations in its temperature are also associated with seasonal changes in the height of the boundary of the troposphere, as air heats mostly from below, near the ground, and in winter the ground is colder and thus heats the atmosphere up to a smaller height.

Quite unexpectedly, meteoric studies revealed another temperature elevation with height, or, as it is called, the upper temperature inversion in the stratosphere. An aeronaut, if he can reach a height of 40 kilometres, will in his fur overcoat have more difficulties in protecting himself from the heat that would there set in after the bitter frost of lower down.

The evidence for the upper temperature inversion comes from the meteoric studies using pictures taken with a shutter. Braking is retarded in the region where the temperature is assumed to increase, as would be expected. In recent years a temperature of +50°C has also been found by direct measurements with the use of probes. It is also of interest that the rate at which the glowing meteor trails dissipate is related to the pressure and temperature of the surrounding air which enables them to be estimated.

It was once thought that the stratosphere was a region of undisturbed quiet, frozen into the stillness of the air ocean, and all the winds and air movements were thought to exist in the troposphere. It, therefore, came as a surprise when the Soviet scientists I. S. Astapovich, V. V. Fedynsky and co-workers discovered air currents at altitudes of 80 kilometres with speeds up to 120 metres per second, which carry meteor trails mainly to the east, but sometimes in other directions; vertical streams are not uncommon.

Meteoric studies in connection with the properties of the stratosphere have just begun, and the findings mentioned above are just their first output, which can convince even stark sceptics of the usefulness of this branch of astronomy.

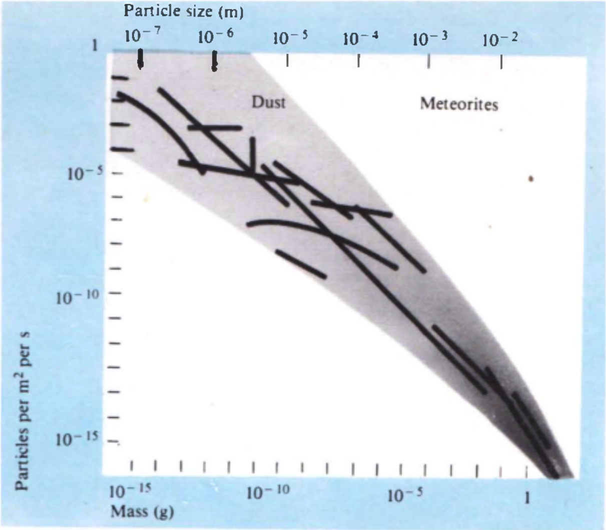
New Techniques in Meteoric Studies

A number of electrically conducting layers consisting of ionized air molecules have been found in the terrestrial atmosphere. The role of these layers is significant as they serve as a mirror for radio waves, reflecting them back downwards, thus making possible radio communication round the globe. Owing to multiple reflection, the waves can travel round the globe between the ground and the electrically conducting layers.

Reflection of radio waves by the ionosphere allows the height of it to be determined. It appears that the ionosphere has different height beginning with 50 kilometres with an ionization maximum at 250-300 kilometres and owes its existence to the ionizing action of ultraviolet sunlight and corpuscules ejected from the solar surface. Changes in solar radiation due to flares on its surface also result in variations in the height and thickness of the ionosphere.

Also, meteoric particles contribute to air ionization. As they are heated and vaporized by friction, meteoric bodies produce the impression of "shooting star" and collide with air molecules to ionize them and themselves. The tracks of these particles last from a split second to several minutes.

On the night of 9-10 October 1946 many astronomers were waiting for the meteor shower, the Draconids, which had appeared in the sky over Leningrad on these days in 1933. These meteors were the fragments of the nucleus of Comet Giacobini-Zinner with a period of 6.5 years. For the first time the swarm encountered the Earth as early as 1926, but then the meteors were scanty. On 9-10 October the Earth approached the cometary orbit with the debris spread along it. On this occasion the Earth was to encounter the meteors that had lagged 230,000,000 kilometres behind their nucleus, i.e. they were closer to the nucleus than the meteors of 1933.

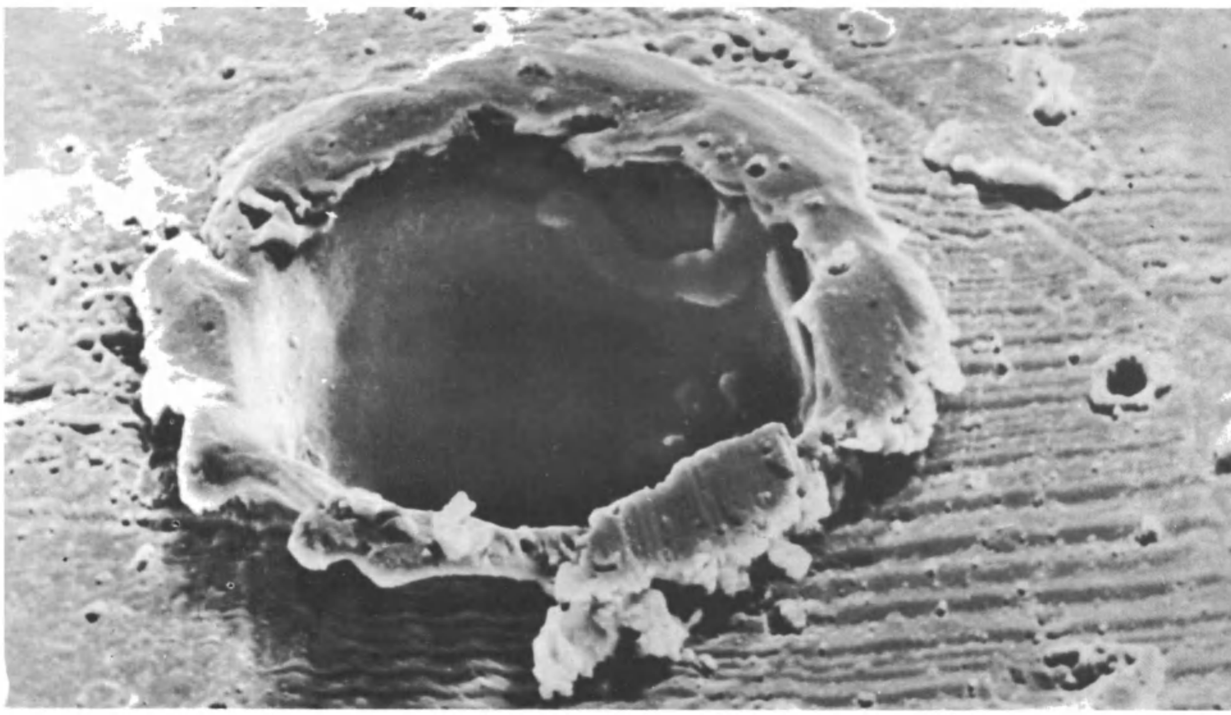


Meteor showers near the Earth. The largest debris are called meteors, the smallest are called dust.

But as ill luck would have it, bright moonlight interfered with the observations making the fainter meteors impossible to see. Few meteors were seen in the evening, the rate increasing slowly towards morning. Their maximum seems to have occurred during the day when they were not seen against the clear morning sky.

But they were "seen" by radars, which on 9 October 1946 helped the meteor hunters for the first time. Radars beamed electromagnetic radiation in all directions and registered each and every reflection of the radiation from ionized meteoric tracks. The weather was clear nearly everywhere, but now it was of no consequence, since even if the sky had been overcast and it had been raining cats and dogs, the radars would still have been up to the mark, for clouds are as transparent to the radio waves they send out as a clean window pane is to light.

So the radar told us that at 4.10 CET the rate was 150 meteors per minute, the maximum. By 7.00 the



Impact crater (0.03 millimetre in diameter) of an iron micrometeorite (*Apollo 16*). The velocity of impact was more than 10 km/s.

shower dwindled away, so indicating that the stream was extremely thin. The Earth took only 2-3 hours to traverse it.

Echo sounding of meteors is now finding ever increasing use. A several thousand kilowatt radar produces directed beams rotated through 360° in azimuth. A radio wave is reflected by a meteoric track back towards the receiver that takes note of the time taken by the wave to come back. The distance from the target to the observer varies, so does the time the echo takes to travel from different parts of the meteor path.

Artificial satellites and space probes are now supplying valuable information about interplanetary debris. We can count the impacts of meteoroids on our spacecraft, determine the rate of their encounters with the craft, their sizes, masses and penetration.

Even the tiniest of pieces of interplanetary rubble can damage a spaceship by depressurizing it, destroying the apparatus and killing the people in it. It was, however, found that the meteoric hazard is not as serious as was once feared. Satellites and probes would send their radio signals to Earth for a long time without being damaged by meteoric impacts.

The instruments in orbit used different methods. Some of them accumulated the energy of impacts of meteoroids and returned the information to Earth. Others counted either each impact or their rate, and so forth. As it had been expected, the smaller the meteors, the higher the rate.

Automatic stations would sometimes encounter streams of meteoric bodies orbiting the Sun. Their number in a unit volume would vary with time. Every thousand seconds each square metre was hit by two particles with an average mass of about $5 \cdot 10^{-9}$ grammes, larger particles were five times less frequent. It was only once that the rate increased 10,000 times.

These numerous impacts were recorded by sensitive instruments, but the impacts were so weak that they did not damage the craft. As to large meteoric bodies, it appeared that there were no encounters with them, and so the hazard due to them was not high, although, it is quite possible that the Soviet probe sent to Venus

in 1962 stopped transmitting before its time because of a collision with some meteoroid.

Until fairly recently, the energy and mass of meteoric bodies had to be deduced theoretically from their observed speeds and brightnesses and the results were uncertain and contradictory. Even large meteoroids escaped observation beyond the terrestrial atmosphere. Now they are taken care of by space stations.

The US *Vanguard 3* satellite registered on 16-18 November 1959 a sharp increase in the meteoric rate, sometimes reaching 200 impacts during a six minute interval, although during one of its two-hour turns not a single impact was noted. This strongly suggests that the meteoric bodies involved belonging to the yearly swarm of Leonids seem to have scattered but slightly away from the orbit of their parent comet. All in all, the satellite had registered 2,800 meteoric impacts in three days of operation, nearly as many as during its other 75 days in orbit. The particles had a relative velocity of 70 kilometres per second, the density of ice, and diameters of about 7 micrometres. All of these particles were smaller in size and mass than those seen as shooting stars in the atmosphere with the naked eye or even through the telescope. The surface of the Moon, being devoid of atmosphere, offers a better place to study the size and speed distributions of meteoric bodies in space. Their encounters with the lunar surface have done no harm to astronauts who have stayed there for a long time.

5. Heavenly Rocks and Dust

One of the sessions of the Paris Academy of Sciences in 1790 was especially amuzing, and the Academicians at the session laughed for a long time. Think of it! The municipality of the city of Juillac in Gascony sent the Academy a report that on 24 July at 9 o'clock in the evening a large rock fell from the sky. It would be understandable if the mayor alone – some sort of lunatic apparently – had signed this nonsense, but the report was also signed by 300 naive residents of the city! "Well, the people of Cascony are known throughout France as innate braggarts", the Academy decided, and at the close of the session they sent a resolution expressing sorrow that the population of Juillac had such a stupid mayor and suggesting that in the future the battle against such superstitions should be carried out more energetically. Actually, materialism successfully and energetically battled against religious superstition and the mystical aspects of the sky in the 18th century. So to what if not sorcery and ignorance could the absurd tales about rocks falling from the sky have been attributed?

Even the prominent French chemist Antoine-Laurent Lavoisier (1743-1794) agreed with his colleagues that "the falling of rocks from the sky is physically impossible".

True, there were cases where such ocurrences had to be acknowledged as fact. So the episcopal consistorium composed a report to the effect that on 26 May 1751 two pieces of iron fell in Hraschina (now Yugoslavia) – the first report on record of such an occurrence.

Russian scientists were among the first to pay attention to this issue and began very early to study rocks from the skies scientifically collecting them and examining their structure and the circumstances in which they fell.

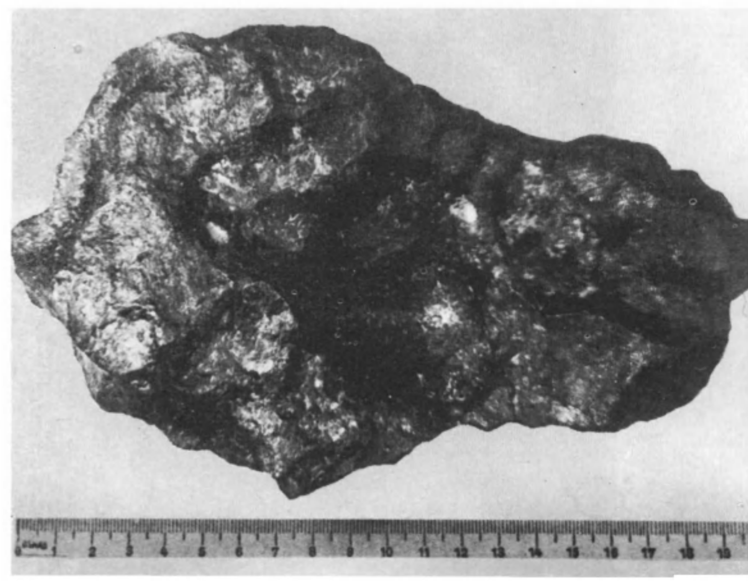
The St. Petersburg Academician Peter Pallas (1741-1811), travelling in Siberia, found in Krasnoyarsk in 1772 a surprising rock, in which iron and stone were interlaced in odd combinations and which local inhabitants considered to be a holy object fallen from heaven. In 1794, when the German physicist and then Corresponding Member of the St. Petersburg Academy of Sciences, Ernst Chladni (1756-1827), learned of the find, he bravely stated that it was possible that the stone had fallen from the sky. He wrote about the Russian meteorite, one of the first to attract the attention of scientists, publishing *The Origin of the Piece of Iron Discovered by Pallas and Some Phenomena of Nature Connected with It*. He proved that such stones do actually fall and can only have cosmic origins. Later he developed the idea and even connected meteorites with comets.

In 1794 a stone fell in Siena (Italy) and a year later another fell in Yorkshire (England) which aided in the conversion of the majority of the scientists of these countries to the views held by Chladni. In France, the very possibility of such phenomena was long doubted. Only in 1803, after a whole shower of meteorites fell in L'Aigle did the Paris Academy send a physicist to the site; the physicist made a report which removed all doubt that such things do happen.

In 1807, Professor A. Stoykovich of Kharkov University published a book with detailed scientific descriptions of all the then known meteorites and a theoretical discussion of their possible origins. Thus the scientific study of meteorites began early in the 19th century but was delayed by the fact that no scientist was an eyewitness to the unusual and unexpected phenomena.

With the dissemination of scientific knowledge, observations and descriptions of meteorite falls became more common. Here is an eyewitness account of a typical meteorite fall that ocurred in 1930.

"On 20 April 1930, the inhabitants of the village of Staroye Boriskino (near the Volga) noticed a small round fireball somewhat smaller than the Moon, flying through the sky at about one o'clock in the afternoon local time, about twenty degrees above the horizon. A string of fire seemed to trail after the fireball. The flight lasted about five seconds, and after the fireball disappeared, a cloud of smoke rose and thickened near the spot where it disappeared, being visible for about five minutes. Soon after the smoke dissipated, a loud report like that of a cannon shot was heard from the west. Then a rumbling was heard, and about three seconds after the report, a second was heard, then a third; in all, there were about ten explosions, about three seconds apart. At first, the reports became stronger, then grew weaker; they seemed to move from west to east, the last report being heard from the spot where the fireball and smoke disappeared. The rumbling after the last report lasted five seconds and gradually died away. About 25 or 30 seconds after the rumbling had died away, they once more heard a sound, at first very quiet, then growing louder, like the wind; it was accompanied by a sound like a rattle (un-



The iron meteorite of Henbury, Australia.

even) and was something like the sound of shrapnel. This sound lasted about 20 or 25 seconds; something seemed to crash – to fall; a sound was heard which was something like 'ooookh'."

No quaking or trembling of the ground was noticed by the eyewitnesses, but they clearly felt that something had fallen to the ground not far away; they thought it had been a bomb. Together with the adults and children who ran after them, they ran into the garden, about fifty people in all, looking for the bomb. Twelve metres from the yard they noticed a dark, round spot about one half metre in diameter. Since there had been no rain for some time, the ploughed black earth of the garden had dried and turned grey, and on this grey background the dark spot of loose, moist earth was clearly seen. There was no depression in the dark spot, the spot being level with the land around. One of the spectators stepped forward, followed by others, and began to dig. The dark soil was loose, and at a depth of ten or twelve centimetres he felt a solid object. He tried to remove it (pull out with his fingers) but the earth was too solid and the object did not come out. Only after he had used a shovel was he able to remove the object. All those present noted that it was a stone, not a bomb or a piece of shell, it was a meteorite. The witnesses said it was about the size and shape of a sheep's head. The meteorite was taken "warm" from the earth, but it was cool enough to be held by hand, no more than 20 minutes after it fell. None of those present noticed anything damaged or burned in the vicinity of the meteorite.

The meteorite was melted on all sides and covered with a dark shell. No one noticed any cracked areas on the meteorite; when it was extracted from the ground, the meteorite was clean, the earth did not cling to it, a smell of smoke came from it.

There are many notable instances of meteorite falls, but we will only discuss a few of them, including some which became generally known only some time after their occurrence.

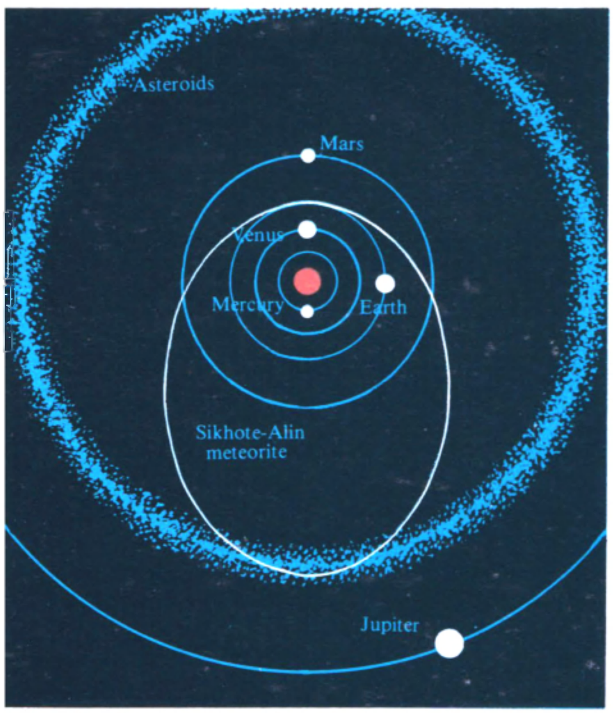
For example, meteorite falls were noted even before Christ in Chinese records. Cases of falls are known from ancient times. So the famous Roman naturalist Pliny the Elder (A.D. 23-79) described the fall of an iron meteorite in A.D. 79 near Naples.

An unexpected difficulty befell the armourers of the Sultan of Khorasan in 1009 when their master ordered them to make a sabre from the iron of the meteorite that had fallen onto this kingdom. The heated iron could not be forged, which is typical of meteorites – the metal is only ductile when cold, and brittle when hot. However, in 1621 Jahangir, Mogul emperor of India, did manage to have two sabres, a dagger and a lance tip made of meteorite iron.

The Eskimos in Greenland, having no concept of iron ore, were able to make themselves knives from meteorite iron they had found in the icy wastes of the Arctic.

More than once, meteorites have been thought of as gifts from heaven and have been made the objects of religious cults. Emperor Heliogabalus built a special shrine for one such holy stone in Rome. The Black Stone of Mecca, embedded in the wall of a shrine, which is the holiest object to Moslems, attracting many pilgrims, is apparently a meteorite, and not a very large one. Meteorites were also placed in shrines in Japan and Alsace (in the 15th century).

The meteorite that fell on 7 November 1492 near the city of Ensisheim on the Upper Rhine was fastened to the church wall with chains by the parishioners, so that the gift from the heavens would not be taken back and would not return to whence it had been sent. The sign placed on the stone by the people of the time is



The orbit of the Sikhote-Alin meteorite, USSR.

remarkable for its expressiveness: "Concerning this stone, many know much, everyone knows something, but no one knows enough." Even in our day, similar signs would be accurate for the meteorites in our museums.

Three descriptions of meteorites contained among others in old Russian writings are especially interesting.

In 1290 in the Northern Dvina area near Veliki Ustyug the familiar picture of a meteorite fall was described in these words: "There was great thunder and the city of Ustyug was terror stricken. People could not hear what they were saying to one another; heaven and Earth shook continuously with a fearful tremor... Countless rocks destroyed the forest and many trees were uprooted while others were completely flattened." This was an entire shower, or more accurately a stone hail, which does not occur often.

Many exaggerations and even plain fabrications are included in accounts of meteorite falls, especially when tales are passed on by word of mouth and not from eyewitness accounts, the more so with delay, when the memory is no longer fresh. A scientist gathering material about a meteorite fall must therefore always keep in mind the Andersen's tale about the chicken who thought the sky was falling.

We must include the tales of fires being started and people being killed by meteorites among such fantasies, although the latter, as a rare exception, could happen. In any case, there are almost no reliable reports of such cases. Cases are known only of meteorites falling on animals (in 1836, sheep were killed in Brasil, and in 1911 near Nakhla in Egypt, a meteorite killed a dog), and of meteorites hitting buildings: in 1684 a meteorite punctured the cupola of a church in Tobolsk in Russia, and in the 18th century a meteorite hit the tower of a Bavarian monastery in Würzburg destroying it. Two or three other similar cases could be cited – no more. It is interesting that sometimes meteorites are found to have fallen very slowly, and therefore to have had no destructive effect. Thus, for example, to the surprise of a laundress, a meteorite fell right into her wash tub, and another, which fell in 1927 and weighed only a few grammes, wound up in the dress of a Japanese girl.

The Stones in Figures

We have already noted that many meteorites have been found long after they fell. By now about 150 meteorites have been collected in the USSR. Most of them are being kept at the Academy of Sciences in Moscow. One of the greats of the collection is the "Pallas iron", which weighed initially 700 kilogrammes. Parts of this stone, which have played a large role in the history of science, have been sent to foreign museums, and now only 514 kilogrammes remain of the meteorite. The second largest in this collection is a meteorite found in 1937 near Minsk. It is much smaller, weighing 188 kilogrammes and is of the same iron-stone material. The largest stone meteorite fell in 1918 near Saratov, it is still in there. The pieces that made up this stone fell along a trail 120 kilometres long and have a total weight of 221 kilogrammes, but the largest single piece weighs about 130 kilogrammes.

The largest known meteorite was excavated in Hoba in south-west Africa where it is still located, since it weighs 60 tonnes and would be difficult to move. It is iron and is unusually rich in nickel (16 per cent) and

difficult to cut. It required two days of hard labour with hacksaws, constantly changing blades, to separate a 2.5-kilogramme piece for chemical analysis. The dimensions of the meteorite are 2.5 × 2.5 × 2 metres.

The world's second largest is the iron meteorite Ahnighito (the Cape York meteorite), weighing 33 tonnes. It was collected in Greenland by the famous explorer Robert Peary in 1897 and taken to New York, but it was discovered back in 1815 and the Eskimos knew of it even earlier.

The third largest is the Bacubirito iron meteorite, weighing 24.5 tonnes, now located where it fell in Mexico.

The fourth largest meteorite, the Willamette in the USA, weighing 14 tonnes, is very interesting. The atmosphere has worn and eroded part of it away; before that, it weighed 25 tonnes.

After these four there is a long list of meteorites, mostly iron with an occasional iron-stony meteorite. The largest of the pure stone meteorites, which fell in Furnas County (Nebraska, USA) on 18 February 1948 is far down the list, since it weighs "only" about a tonne.

Sometimes several pieces fall at once, apparently the result of the destruction of one body in the Earth's atmosphere. Sometimes the parts can be fitted flush together, but sometimes the parts seem to have been separate as they flew through space together. We have already mentioned entire showers of stones. These are apparently meteorite clusters, and they sometimes fall over areas hundreds of square kilometres.

The table below contains a list of the meteorite showers of the last two centuries that have been studied by scientists.

Meteorite Showers

Location	Date	Number of fragments
L'Aigle (France)	26 April 1803	3,000
Knyahinya (Czechoslovakia)	9 June 1886	1,000
Pultusk (Poland)	30 January 1868	3,000
Hässle (Sweden)	1 January 1869	1,000
Homestead, Iowa (USA)	13 February 1875	100
Mog (Rumania)	3 February 1882	3,000
Holbrook, Arizona (USA)	19 July 1912	14,000
Saratov (USSR)	6 September 1918	200
Pervomayskaya (USSR)	26 December 1933	100
Sikhote-Alin (USSR)	12 February 1947	Many thousands

If meteorites are counted not by the piece, but by "falls", it turns out that each day the Earth receives 5 or 6 (2,000 per year), but only 2 or 3 per year are observed, sometimes not even one, since the great majority of them fall in uninhabited areas. The vast Arctic and the oceans occupy a large part of the Earth's surface; they are inhabited by polar bears and fish who are known for their silence and are therefore frequent though useless witnesses to meteorite falls. In the deserts, steppes and forests, there are not very many people, and even if a fireball is seen, it is often not easy to find the meteorite. At night, even meteorites which fall in populated areas may go unnoticed. That is why, of the 200 tonnes of meteorites which "land" each year, we find on average each year only a few kilogrammes. This estimate is approximate, naturally, and some consider that the yearly meteorite fall is closer to 2,000 tonnes.

The addition to the Earth's mass is mainly in the form of stone, if we accept the average of the meteorites whose fall has been observed. Iron meteorites make up only 5 per cent, and iron-stony, only 1.5 per cent. This is probably the true ratio of stone to iron meteorites throughout the outer space.

Incidentally, of the found meteorites whose date of fall is unknown, 66 per cent are iron. The reason for this difference is naturally that iron meteorites attract more attention, especially in the steppes, the desert or the forest. Often only a specialist can determine whether or not a given stone is of meteorite origin, and only after a special investigation.

If we look at the map of distribution of meteorites, we will see a strange peculiarity: meteorites for some reason tend to fall near culturally inhabited areas, primarily along railway lines. The reason for this, of course, is that where the population is sparse and perhaps less educated, fewer meteorites are noted and those that are found are less likely to be brought into scientific centres and hence registered on the rolls of science.

In the USSR, meteorites are the property of the state, and there is a reward for finding a meteorite and bringing it to the Academy of Sciences. In the capitalist part of the world, the owner of the meteorite is the owner of the land on which it falls. Therefore, meteorites there are objects of commerce, they are bought and sold, and people, who are not interested in selling their meteorites often break them up into pieces from

personal curiosity, or bore holes in them, thus destroying their value to science. Jules Verne, in his novel *The Golden Meteorite* masterfully portrayed the agitation resulting in the bourgeois society from the imaginary fall of such a meteorite. No golden meteorite has ever fallen to the Eath, and it is highly unlikely that such a meteorite has ever existed. Although diamonds were found in the Novo-Ureysk meteorite, they were microscopic. They were very valuable scientifically but, due to their extremely small size, they had no market value, and it will do no good for a vain girl to dream of ear-rings containing diamonds which fell from the sky. We will discuss the chemical composition of meteorites below, and now we will speak a little of the outward appearance, by which it is possible to tell a meteorite from an earthly stone. Unfortunately, few people know these signs, and the author has more than once wasted time travelling to investigate something to satisfy the praiseworthy enthusiasm of a local resident, only to find an ice-age stone or perhaps even a cobblestone.

Structure and Age of Meteorites

Iron meteorites, as was said above, are easier to find and dig up, and are rust-brown in colour. They are always irregular, and the surface, if not oxidized, is covered with a thin, smooth, black crust called scale. This crust is a result of the melting of the outer layer of the meteorite as it falls through the air. The meteorite falls so rapidly, however, that at a certain size, the centre of the meteorite does not burn, and the melting of the surface is confined to a thin layer at the last stage of its (decelerated) fall through the atmosphere. The temperature of a meteorite in fall and flight is almost the same as when it flies by the Earth. This is the temperature of a body in space, heated by the Sun at the distance of the Earth. This temperature is about 4°C. Contrary to science fiction stories the centre of a meteorite is neither red hot nor cooled to absolute zero (−273°C).

When polished and etched in weak acid, the surface of meteorite iron becomes covered with a tracing like that on a frozen window, caused by the properties of the crystal structure of the iron. This tracing is called the Widmanstätten figures and is a certain means of identifying meteorite iron from natural or ore-smelted iron.

Stone meteorites are also usually covered with a thin, black, glassy crust, sometimes rough, sometimes smooth. It is eroded and oxidized if the meteorite remains long in the open air or in the ground. It is then even harder to tell a meteorite from a terrestrial stone. The inside of meteorites, when broken open, is different. Most often it is grey, sometimes with round grains of a peculiar structure (called chondrules) and with metallic particles.

The polished surface of a meteorite, observed under the microscope, presents for the specialist a peculiar structure, differentiating it from a stone of earthly origin even though both have similar chemical and mineralogical structures. This specialist is not an astronomer, but a mineralogist, or more correctly a pertographer* and one who specializes in the study of meteorites. Under the supervision of Academicians Vernadsky and Fersman, a whole school of such specialists has grown up in the USSR: P. L. Dravert, P. N. Chirvinsky, L. A. Kulik, and others. The meteorite is in the province of the astronomer only so long as it is a heavenly body, i. e. is not on the Earth. The astronomer could also meet such a guest on the threshold of his home, the Earth, i. e. he can determine its trajectory in the atmosphere, but to delve into the details of the structure of the stone requires another special education and great experience in the study of stones and minerals. The science of petrography, at the stage of detailed study of meteorites, divides them by structure into a multitude of classes separated by various properties.

When a meteorite is in the atmosphere, a strong "wind" blows around it from the front and sides and, melting the surface, removing from it first the easily melted materials, and also in general smoothing down sharp edges and angles. Therefore, a characteristic feature of a meteorite, if it was not broken in the final moment of its journey, is that its lines are smoother and more curved than when it was in space. The air, so to speak, machines the meteorite, but the result of this treatment depends on the velocity of the meteorite, on its form, and on its turning in flight. Often, meteorites resemble a piece of clay shaped by hand. Its surface has furrows, depressions, sometimes even cracks, running in all directions from the front of the meteorite.

* From the Greek word *petros* for stone.

The meteorite itself is conical in shape, like the nose of a shell.

The detailed chemical composition of meteorites will be described in the next paragraph. I. Mukhin at St. Petersburg analysed meteorites chemically before 1819. Recently, not only qualitative, but also quantitative chemical analyses of meteorites have been performed in great detail. Alas! This necessary curiosity was very expensive for us, since for this chemical analysis it was necessary to destroy, literally pound to powder, a large number of meteorites from museum collections. These meteorites can now never be scientifically investigated further, and the meteorite investigators who are not chemists cry, "Enough of chemical analysis, we are satisfied with what we know about the chemistry of meteorites! Leave us someting so we can study the dimensions, form and structure of the meteorites."

We have already presented the average chemical composition of stony meteorites, which varies somewhat from meteorite to meteorite. Basically they consist of oxygen (36.3 per cent by weight), iron (25.6 per cent), silicon (18.0 per cent) and magnesium (14.2 per cent). The other chemical elements (not all of those which are known on Earth) are present in amounts of one per cent or less. In general, they are similar in composition to the Earth's crust, especially the deeper rocks. In comparison, the earthly rocks contain more silicon and oxygen, and less magnesium and iron. In place of the latter terrestrial minerals contain aluminium, but, apparently, the deeper we go into the Earth, the closer the composition of the Earth's strata approaches that of the meteorites.

Iron meteorites, in addition to iron (91 per cent) and nickel (8 per cent), also contain cobalt (0.7 per cent), phosphorus (0.2 per cent), and trace quantities of sulphur, carbon, chrome, and copper.

Only up to 0.0004 per cent of gold has been found, i. e. if all the gold could be extracted from all the meteorites so far collected on Earth, there would be less than one kilogramme. However, even this is practically impossible, since the gold in meteorites is powdered; it would be like trying to make a living selling the pins lost by hikers and found among the autumn leaves in the forest.

It is interesting that in 1946, the Soviet petrographer L. G. Kvasha found 8 per cent water in one meteorite, although it was bound in the minerals, not free.

The iron meteorite of Cape York, Greenland, with Widmanstätten figures.

The radioactive elements are present in meteorites to an even lesser degree than gold; radium accounts for 0.00000000001 per cent, or 20 times less than there is in rocks on Earth. However, finding this tiny quantity of radioactive elements in meteorites is incomparably more important than finding gold or diamonds, even if the latter were a hundred times more plentiful than they are.

The radioactive elements and their partner – helium gas – are sort of "birth certificate" for meteorites, since they tell us the age of our heavenly guests.

Uranium and thorium are known to decompose spontaneously into other chemical elements, giving off heat, electrons, X-rays and atoms of helium. At the end of this chain of atomic transmutations is lead, which has no tendency to decay further. The "obstinacy" with which the atoms of radioactive elements decay and follow the law of this decay, ignoring all attempts to accelerate or retard the process, is also known.

No matter how much uranium there was to start with, after 4,560 million years one half of its atoms will have decomposed, that is after 4,560 million years, 0.5 gramme will remain of each gramme of uranium. After another 4,560 million years, one half of this half gramme will remain, i. e. 0.25 gramme. The same is true of thorium, although it decomposes more slowly,

losing one half in 13,000 million years, whereas radium (an intermediate product in the decay of uranium) decomposes much faster: after only 1,600 years it loses half its mass.

The light helium atoms given off by the nuclei of the heavy atoms of the radioactive elements are accumulated in the solid mass which contains them. It is not difficult to determine how much helium should be accumulated by the decay of, say, 1 gramme of uranium. It is therefore not hard to determine how long the uranium has been decaying in a given rock containing a known quantity of uranium and a known quantity of helium. Obviously, the uranium and thorium in a rock have been decaying for the length of time that they have been in the rock, i. e. since the rock formed, say, solidified from a molten mass, after which the helium could no longer evaporate and the uranium could no longer escape. After solidification, the uranium and the decay products are confined to a life sentence, as in a jail.

Thus, the ratio of helium and uranium in a rock determines the age of the rock, and with greater accuracy than that with which we can assess the age of a man by looking at him.

This method has been used to determine the age of rocks, and it has been found that the oldest rocks in the Earth's crust are 3-3.5 milliard years old. This is then the approximate age of the Earth's crust, and a very venerable age it is.

Panet and associates performed the very difficult determination of the content of uranium and helium in many meteorites–difficult because there is so little present. The results obtained for several dozen meteorites lead to an unexpected conclusion.

It turned out that the "ages" of the meteorites ranged from 60 million to 7,600 million years. It seemed that the scientists had in their possession some very young interplanetary bodies, since for a heavenly body, 60 million is very young indeed.

But it soon transpired that the seeming difference in age was not a genuine difference in "lifetime", but rather in "living conditions". The fact is that the ratio of helium to lead in a meteorite depends not only on its age, but also on the treatment it receives from cosmic rays–particle streams with great energy. But when the helium of "internal" origin was separated from the helium of "cosmic" origin, which appeared to be quite a problem in itself, the age of the meteorite was found to be more realistic: from 2.5 to 4 milliard years.

Incidentally, we have not yet discussed the mineralogical and petrographic structure of the arrivals from the sky. Actually, the same atoms can enter various molecules, producing various combinations, and what is more, they can form more complex compounds, called minerals.

The main minerals making up stony meteorites are known and occur widely on Earth: olivine, pyroxene, feldspar, plagioclase, and nickel iron. Many terrestrial minerals, however, are not found in meteorites, for example orthoclase and mica, although they are very common on Earth.

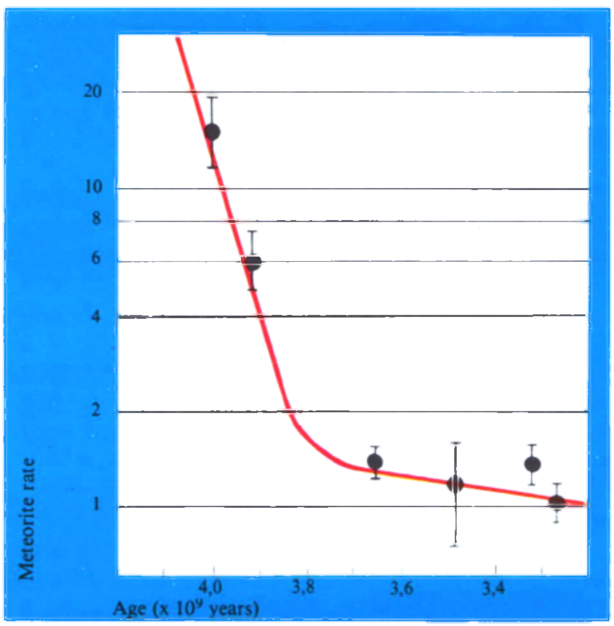
Also, meteorites have acquainted us with some minerals that for some reason or other are not formed on the Earth. They are named after the scientist who discovered them: schreibersite, daubreelite, moissanite, etc.

The chemical and mineralogical composition studies of meteorites support the very important philosophical conclusion of the material unity of the Universe. Beyond the limits of the Earth we encounter, for example, the same chemical elements which the great Mendeleyev placed in his table, and the elements which have since been added to the table. The laws of chemistry are thus seen to be true not only on the planet where they have been established. But at the same time, nature does not display the tiring uniformity which some metaphysical minds have tried to attribute to it. The mineralogical variations in meteorites and the presence in them of minerals not found on Earth are clear examples of the multiformity of Nature, which is caused by the infinite qualitative forms of movements, processes which take place in an eternally existing and ever-changing Universe.

The Chemical Composition of Earth and Meteorites

We know that in meteorites no chemical elements are found that are not found on Earth, but to what degree are the abundances of the chemical elements found on Earth and in meteorites similar?

Unfortunately, it is difficult to analyse the composition of the Earth, whose core is not open to us. Even during volcanic eruptions, comparatively shallow-lying material is brought to the surface of the Earth. The composition of the Earth's core can only be



Rate of occurrence of meteorites as a function of age from the investigation of Moon rock samples.

judged directly, by studying compression inside the Earth, the slight periodic changes in its form under the influence of the gravity of the Moon and the Sun and mainly by studying the propagation through its mass of the vibrations produced by earthquakes. This indicates the hardness and elasticity of the Earth's interior and the variation of density in it. Also, it has been discovered that material from deeper layers inside the Earth is richer in iron than material from shallower layers.

The sum total of these data and various other considerations have led scientists to the conclusion that the Earth has a dense iron-rich core, surrounded by a rocky shell, solid only on the outside, the thickness of the crust not exceeding 200 kilometres.

Our Earth is, if you like, a sort of "cherry", with a solid core-"stone", and a rocky "pulp" which really is a pulp, it is plastic, and even the core, although very dense, is still not solid in the normal sense. The material inside the Earth is in a very special state, subject to high temperatures and terrific pressures. We have not been able as yet to create these conditions in the laboratory and transform material to this special state.

As a whole, the Earth reacts to rapid external influences like a steel ball, and to slow, but prolonged, influences like a piece of clay, or a balloon full of water. The Earth thus combines elastic and plastic properties.

When we work out the average chemical composition of the Earth, we must keep in mind that the composition of the Earth's crust, with an average density of around 2.7 grammes per cubic centimetre is not the same as that of the entire planet, which has an average density of 5.5 grammes per cubic centimetre. The larger planetary figure is a result of the increase in density with depth, which reaches 11 in the solid core.

The average chemical composition of the Earth, as determined by Fersman, is shown in the table in comparison with the average composition of meteorites.

Average Abundance of Elements (Weight Per Cent)

Elements	Earth	Meteorites	Elements	Earth	Meteorites
Iron	37.6	38.0	Nickel	3.0	2.8
Oxygen	29.0	29.0	Calcium	2.0	1.1
Silicon	14.5	14.4	Sulphur	1.5	1.9
Magnesium	9.2	11.0	Aluminium	1.5	0.6

The remaining chemical elements account for much less than one weight per cent each.

True, the average abundances cited for the Earth are given with due regard to the composition of the meteorites that have fallen on Earth.

These calculations thus assume that the chemical composition of various heavenly bodies should be similar, but, as we shall now see, this table takes account not only of meteorites. Here, for example, are two more "pure" comparisons.

The table is based on the supposition that for meteorites the stone-to-iron ratio is 4 to 1, since that is about the current ratio. So an average meteorite contains 38 per cent iron. The Earth's core, which can be taken to consist of iron on the strength of its elastic behaviour, has a diameter of 3,500 kilometres, this figure having been established to within reasonable accuracy. With a density of 11 grammes per cubic centimetre, this gives about $^1/_3$ of the total mass of the Earth, i. e. we see that without any assumptions, the main chemical element–iron–is as abundant in the Earth as it is in meteorites.

Incidentally, it has recently been concluded that the iron abundance of both the Earth and the meteorites is lower than that indicated and that the great density of the Earth's core is mainly due to tremendous pressures rather than to a predominance of iron.

In the stony terrestrial core itself some increase in the percentage of iron would be expected because of the crust's sinking, which is indeed observed. If we do not consider the less studied deeper rocks, the so-called *peridotite* rocks, and confine our discussion to granite rocks containing the least percentage of iron (the upper portion of the Earth's crust – about 20 kilometres thick) and the plateau-basalt rocks (a layer of 40 kilometres below the granite layer) we obtain the table.

Average Abundance of Elements (Per Cent)

Elements	Stony meteorites	Earth's crust	
		granites	basalt
Oxygen	36.3	46.6	44.3
Iron	25.6	5.0	10.4
Silicon	18.0	27.7	23.9
Magnesium	14.2	2.1	3.8
Aluminium	1.5	8.1	7.0
Nickel	1.4	–	–
Calcium	1.3	3.7	6.9
Sodium	0.6	2.8	1.9

From this table it follows that the similarity between the chemical composition of the Earth's crust and stony meteorites is very great.

The Earth's crust contains radioactive elements, like uranium and thorium, and the heat produced as they decay raises the temperature with depth, which can be observed when descending down a mine shaft. On average, the temperature increases 3° with each 100 metres, although sometimes greater rises are observed. For example, the author descended down a mine during a sharp frost on the surface and was soon sweating with the heat. If this increase continued to the centre of the Earth, we would find there a temperature higher than that on the Sun, which is not the case, as was said earlier. Apparently, the temperature in the centre of the Earth reaches only 2,000°C and falls off toward the surface, more rapidly as the surface is approached because the Earth loses heat into space from it.

The radioactive elements are only present in the crust, the deeper strata having them in smaller amounts. Heat accumulates in the centre of the Earth, and is lost from the surface. It is possible that in general the Earth may be slowly growing warmer.

In Search of Ancestry

The clue to the ancestry of meteorites lies, however, not in their mineralogical structure, but in the laws governing the ways in which these minerals may combine.

It turns out that the chemical elements of which meteorites are made up are the ones that occur on Earth, and that most of the constituent minerals of the meteorites are known to us. Petrographically, however, these meteorites are very different from terrestrial rocks. First of all, the meteorites contain no sedimentary rocks, such as sandstone and limestone, which are formed from layers of sand and shells on the ocean floor being pressed together under huge pressures. As to lime, it is organic in origin consisting of the shells of creatures that once were alive; no traces of any living thing have been found in meteorites. They do not even contain bacteria, which are almost impossible to avoid on Earth.

Most stony meteorites (*chondrites*) contain a multitude of tiny spheres resembling solidified drops of glass and called chondrules.

Three per cent of stony meteorites, otherwise no different structurally, are black in colour. This is, probably, a result of their being severely burnt at some time. Not just on the surface as it generally occurs when they pass through the atmosphere, but clear through.

Laboratory experiments have shown that heating grey meteorites to 800°C for several minutes causes them to turn black. In order to be heated to 800°C by the Sun, the meteorites would have to approach it ten times closer than the Earth does. Consequently, the orbits of many of them, if not of most, should have perihelia within the orbit of the Earth. Unfortunately, unlike meteoric orbits, we know too little about meteorite orbits. If it is found that the orbits of meteorites do not display this characteristic, we will have to assume that these orbits were taken on at some later stage of the development of meteorites, as a result of drastic changes in the orbit caused by planetary disturbances.

It is possible that this heating could have been

caused by another star, if the particular meteorite was not a member of the solar system. Knowing the separation between stars, we can calculate that such an approach could occur once in milliards of milliards of years. However, the meteorites are much younger. Therefore, the supposition is highly unlikely.

Not knowing the exact orbits of meteorites obscures their origin. In this respect, the petrographer and the geologist have, I believe, much more to say than the astronomer, who can only state the following.

If not all, then at least some, meteors are associated with comets, and meteors vary widely in dimensions and mass. In terms of dimensions, meteorites approach meteors, since sometimes fine cosmic dust falls on Earth. It is also hard to believe that a cometary core could only consist of small particles and contain no stones the size of meteorites. But if this is so, then meteor showers, it would seem, should accompany the fall of meteorites. No such connection has been found between the two.

Judging from the spectra of meteors, those that undoubtedly come from comets are stony. Stony meteorites can therefore be attributed to comets, and iron meteorites and meteors to some other source. Again it is difficult to connect meteorites with comets, besides it is still unclear where the iron-stony meteorites, in which the stone mass has the same structure as in purely stone meteorites, come from.

The gradual transition from meteorites containing almost no metal to those rich in it, and through iron-stony meteorites to iron meteorites, causes petrographers to seek for common source for them all. The structure of the stone and the crystallization pattern of the iron in meteorites strongly suggest that they can only be described by the slow cooling from the liquid state.

An opinion some people have held for a long time and which is supported by A. N. Zavaritsky holds that meteorites are the remnants of an object that perished due to a collision, the object being a planet similar in size and mass to ours. Whether this planet was located where the asteroids are today and they are only the largest of the debris, or whether it came from outside the solar system cannot be said for certain, but the former is much more probable. There are various theories concerning the causes of the catastrophe that destroyed the hypothetical planet located at one time between Mars and Jupiter.

The surface cooling of the Earth should have led first to the formation of high-melting materials – olivine and pyroxene – which would be mixed with the remaining, still plastic, mass and even find their way into the iron-rich material. The solidified rocks and the plastic metal should be mixed the most in the intermediate strata. The succession of meteoric structure is too complete to be considered random, and is very much like to variation of stone-to-metal ratios likely to occur with increasing distance from the centre of the cooling Earth. It is only in the depths of the planet that large-size iron crystals could have formed that are peculiar to meteorites. The age of rocks varies depending on their formation time. If an explosion destroyed an Earth-type planet the resultant debris would have a wide variety of sizes and shapes, and differing in iron content and rock age. The largest of these fragments could be asteroids. Recently, some asteroids have been discovered that are very much similar to meteorites in size and eccentricity of orbits that intersect the Earth's orbit. From these, it is possible to go over to the sporadic meteors. Characteristically, it seems that the largest orbits are found in the smaller asteroids, which in an explosion could receive the greatest side thrust and could be hurled into the larger orbits, so that they would be the most influenced by other heavenly bodies. For these two reasons the orbits of some of them should be markedly different from what would have been the orbit of the dead planet that laid between Mars and Jupiter.

If we take the planets from the Earth, through Venus, Mars and Mercury to the Moon, we will observe, along with the decrease in size, a decrease in the average specific density of the planets, reaching 3.3 with the Moon. Since the lighter, exterior parts of the Earth's crust and the stony meteorites both have a density of 2.7, it is obvious that the size of the iron core decreases even more rapidly than the size of a planet, and therefore the former planet could not have been much smaller than the Earth. Otherwise, its core would not have been sufficient for the formation of the host of iron meteorites, which from the viewpoint of mass are the intermediate link between normal meteorites and the smaller asteroids. Let us now look at them in more detail.

Canyon Diablo

Virtually all the features of the ground – mountains, valleys, ravines and lakes – are all the results of either the internal activities of the Earth, or the action of water and air. In general, we would be right in saying that the Earth itself formed its own physiognomy, the face we know so well today was cast in its tempestuous youth. In those times streams of hot lava came from the Earth's bowels more often than now. The untiring work of the winds and water brings down the mountains and erases the traces of this past, although even now at times we are reminded of the convulsions of the Earth's crust, the reminders being destructive earthquakes. At the same time, these earthquakes are creative processes, pushing up new mountains or leading to sinking the land. The formation of a ridge a thousand metres high is accompanied, naturally, by many earthquakes lasting millions of years.

Geologists can explain each of the formations of the Earth's surface, but one.

Meteorite crater Canyon Diablo, Arizona, USA.

Imagine a vast and empty plain, where for millions of years no earthquakes have occurred, no volcanoes have formed, and yet in the centre there is a huge crater. It is not similar to volcanic craters – they are comparatively small depressions at the summits of cone-shaped mountains. Near this crater there are no lava streams, or other trace of volcanic activity.

In Arizona (North America), on a high plain consisting of horizontal layers of limestone and sandstone, there lies, in a wide ring, a stone wave surrounding a depression 1200 metres in diameter. Its floor is 180 metres below the surrounding ridge, and the walls of the stone ring rise 45 metres above the surrounding plain. The huge pit was formed by some sort of grandiose explosion, since the ring consists of thick stone strata of limestone and sandstone, broken and moved aside by some monstrous force. In the southern portion of the crater, a stone wall 500 metres long is even vertical, rising 32 metres above the strata of the same rocks. The ring around the crater, its walls and floor are piled with heaps of stone chunks, including many rusted pieces of iron.

The area and floor of the crater are filled with a layer of "stone flour" – fine lime and sand powder – the result of an impact crushing the rock. It is even more interesting that many stones in the crater show signs of melting due to high temperatures and contain impressed particles of nickel iron. Pieces of rock weighing up to 700 tonnes have been thrown out around the crater up to 10 kilometres away.

The local Navajo Indians called the crater "Canyon Diablo". The legend in the tribe is that some time long ago a flaming god fell into the crater.

Only in 1891 did the canyon attract the attention of scientists, who speculated that it was not the result of normal geological processes, but was rather produced by some catastrophe, perhaps, the fall of a gigantic meteorite. Investigation of the irregular pieces of iron found around the canyon and especially within it confirmed that it was meteorite iron, not only because of their form but also because of the presence of the Widmanstätten figures which appeared on their surface after polishing and etching with acid.

Thousands of pieces of meteorite iron were found up to several kilometres away from the crater. They weighed between several grammes to 500 kilogrammes, many of the smaller ones having already rusted through and turned to powder. By now, collec-

tors have gathered up almost all the rock that could be found on the surface, and even earlier, the Indians collected the largest pieces for various purposes. Now, with the help of electrical and magnetic means like those used to locate ore deposits, pieces buried from 2 to 3 metres beneath the surface are being brought up.

However, neither a giant meteorite nor its remains have been found in the crater. This suggested that the giant meteorite which made the huge crater lay below it, covered by layers of rock.

The location of the giant meteorite which fell here would be of great scientific interest but it would require deep boring, for which no American scientific organization has put forth the money.

For this reason, one of the crater's investigators, a man called Barringer, tried to interest some American financiers in the project on the grounds that the great iron mass of the meteorite would have a commercial value. Pure iron is not found on Earth, and an exploitation of the meteorite could be a profitable venture. Thus, a company was organized to finance the drilling.

The diamond bit chewed its way into the stone in the centre of the crater. At several locations the drilling went down 300 metres and in all cases the boring revealed, first ground and powdered sandstone containing some iron chunks, the proportion of iron petering out the deeper the bore went. Finally, the bore reached solid horizontal layers of white, then red sandstone, which were apparently unaffected by the fall of the meteorite. The meteorite was not found.

Devil's Canyon, having pulled aside the curtain masking the mystery of its origin, let it fall back in place once again...

The failure did not, however, dampen Barringer's spirit. He remembered that artillery shells and bullets, when falling into dust, form round craters, regardless of the angle at which they fall. This suggested that the meteorite might have fallen into the Earth at an angle. Hence it might have got buried not under the centre of the crater, but under one side perhaps under the ring wave formation. It was noted that the raised layers of limestone forming the base of the wall were inclined in the north only at 5°, to the east and west somewhat more steeply, and in the southern portion of the crater the layers were completely vertical. Apparently, the meteorite might have flown north to south at a rather small angle to the horizontal and, causing the greatest destruction in front of it before penetrating the southern wall to be buried there. This conclusion was supported by the first borings. The borings which had been made in the bottom of the southern portion of the crater had to go to great depths to uncover undisturbed layers of sandstone.

The picture revealed by the new borings made in the southern wall of the crater was completely different. At first the drill penetrated inclined but solid layers of sandstone, then it entered a fragmented mass, and, beginning at 385-metre depth, the number of meteorite particles encountered began to increase sharply. At 410 metres, the matter encountered was 75 per cent nickel iron, which greatly resisted the penetration of the drill. A few dozen metres deeper the drill became stuck in this mass, broke, and remains there today. It was decided that the drill had encountered the main mass of the meteorite. That was in 1927.

Unfortunately, further work was curtailed, since the meteorite was found to be at an unexpectedly great depth, and digging it out would have required the construction of a deep mine, the control of the underground waters encountered, and so forth. Such an undertaking would be expensive and would not have been justified by the income from the iron removed, therefore the disappointed financiers, who had formed the exploration company and who were little interested in scientific aspects of the venture, curtailed the funds. The work performed had already cost $300,000.

The gigantic Arizona meteorite, which blasted the huge crater in the Earth reminiscent of lunar craters, seems to have been found, but left buried where it fell. Incidentally, the experts are divided on this difficult question, and, as we shall see further on, the majority believe that the meteorite was destroyed on impact and that nothing lies beneath the wall of the crater.

How long ago did the meteorite fall? We can only guess, by judging from the degree of oxidation of the meteorites that have been found, the degree of erosion of the crater by water and wind, that the great event occurred not less than 5,000 years ago.

Other Meteorite Craters

The success of the investigation of the Arizona meteorite crater spurred the energy and imagination of scientists in all lands, and in the last 25 years a number of other craters, also of meteorite origin, have been

discovered and investigated. They are not as large as the Arizona crater, but studying them greatly increases out understanding of the behaviour of these heavenly strangers which find their final resting place on our planet. They travelled the airless wastes for millions or even milliards of years and, after making bright fireworks in our atmosphere, they created rare formations on the surface of the Earth, changed its relief, and parts of them remain buried beneath the products of the destruction they caused. Since these new data are little known and very interesting, they are worth considering in more detail.

First of all, a crater 162 metres in diameter was found in 1921 14 kilometres outside Odessa, Texas, a town Russian emigrants named after their home city. The crater's border is not very steep and it is only 5 metres deep. The bottom and edge are also formed from pieces of broken sandstone and limestone, and many pieces of meteorite iron were found there. Later, strong magnets were used to separate another 1,500 pieces of iron. The edges of the crater disrupted by the explosion are inclined 30° to the horizontal.

Near the city of Brenham, Kansas, about a tonne of iron meteorites was collected in 1885 in an area of several square kilometres. However, as recently as 1933 a long depression (11 × 17 m) 3 metres deep was noted. Within the crater, several meteorites of up to 25 kilogrammes were found, along with hundreds of partially oxidized pieces.

South America has been less studied geographically than North America, and perhaps for this reason, only one trace of the fall of a large meteorite has been found there.

In Campo del Cielo, Argentina, a group of round depressions, the largest about 70 metres across with a wall of 1 metre high, have been known since 1576. Over a tonne of meteorite material has been found nearby. Not long ago, in one of the craters, with a diameter of 53 metres and a depth of 5 metres, digging was performed, and glassy pieces of melted stone and pieces of meteorite iron like those found in the Arizona crater were located.

In Europe, almost every square inch of which is well known, only one group of meteorite craters has been found; it is in the small island of Saaremaa (Oesel) in the Estonian SSR. Six of these craters, known for over a century, were just recently subjected to scientific investigation. The largest of them was filled with water, the lake measuring 90 × 110 m with banks rising 6 metres above the surrounding land. The sloping walls and smaller craters contain a great deal of rock flour and pieces of stone. The craters are almost round and are 35, 33, 20 and 10 metres in diameter. The sixth crater is oval (53 × 36 m) and is formed, apparently, of two adjacent round depressions. The walls of all the craters are formed of inverted and broken layers of limestone, inclined from the centre of the sides at angles of 30 to 40 degrees. On the bottom of the smallest of the craters, after the pieces had been removed, a depression 0.5 metre deep with melted sides was found. In 1937, after careful digging and research, 110 grammes of meteorite iron comprising 28 pieces was found. This resolved any doubt as to the meteorite origin of the craters. The small weight of the meteorite material found on the site is explained apparently by the fact that the local population of the island had already removed all the conspicuous meteorite material. Thus whilst the existence of population in an area aids in the location of meteorite craters, at the same time it destroys traces of the event. If the craters had been in some deserted location, they would have remained unknown, but the traces of the fall would have been preserved with greater integrity and science would have been able to produce a more accurate and more complete story of the formation of these unusual features. A meteorite scientist is often like a detective, to whom it is important that the scene of a crime that occurred without witnesses be undisturbed.

It is not surprising therefore that the majority of meteorite craters known today are in empty places. Such, for example, are the Arabian and Australian craters.

In the great Arabian desert Rub al-Khali, where in ancient times roamed the militant bedouins, a group of craters lies. It was found in the search for the legendary city of Wabar, which had been destroyed by "heavenly fire" for the sins of its inhabitants. The walls of the "city" were found to be the walls of meteorite craters, two of which are still visible. The rest are probably buried beneath the shifting desert sands. The largest has a diameter of about a hundred metres, a depth of about 12 metres, and the other is about 40 × 55 metres, although they too are partially filled with sand.

Iron meteorites were found near the craters – the largest weighing 11 kilogrammes and glassy masses,

made of sand melted when the meteorite landed. Particles of nickel iron are included in these glassy masses.

We can get an idea of the high temperatures produced in a meteorite fall if we remember that iron melts at 1535°C and silicon at 1710°C. Thus, the presence of glassy particles and stone in meteorite craters indicates that when meteorites hit the ground, their energy is partially transformed to heat and the temperature rises to 1500°C. Naturally, this temperature only occurs for a very short time, and the melted particles solidify again almost immediately.

The Australian craters in the region of Henbury were discovered in 1931. In an area of 1.25 square kilometres were discovered 13 craters with dimensions to 200 × 110 metres, the rest being round and having diameters from 9 to 18 metres. The largest has a maximum depth of 15 metres. In the smallest, 4 pieces of a meteorite, having a total weight of 200 kilogrammes, were found 3 metres below the floor of the crater. Another 800 or more pieces of meteorites were found in the area.

In these craters as in the others we find rock flour, distorted and intertwined particles of rock and melted sand in the form of glass droplets and fibres thrown out up to 1.5 kilometres away.

The local Aborigines are afraid of the area, and they call it the "devil stone flying with the Sun's fire". This name, like the legend of the Indians about Devil's Canyon and the story of the city of Wabar, recalls the flight of the fireball. Incidentally, these craters were formed long ago, probably thousands of years ago. This is an interesting example of the length of time legends can be retained and passed on.

Three hundred kilometres from Henbury in Box Hole, another crater 175 metres in diameter was discovered at a later date. It is 10 to 16 metres deep and its sides rise 5 metres above the surrounding ground. It has been greatly eroded by the atmosphere and is probably very ancient. In 1937 an iron meteorite weighing 82 kilogrammes together with rusted fragments were found there.

A 70-metre crater 5 metres deep was located near Dalgaranga in Australia in 1923. Pieces of bizarre shaped meteorite iron were found near it. Also, a 840-metre crater named Wolf Creek was found in Australia in 1948. Its rim rises 30 metres above the surrounding territory, and its floor is equally deep.

In 1946, the world's largest meteorite crater was found in Canada, the Chubb crater (or, as Canadians now prefer to call it, New Quebec crater). It is about 3.5 kilometres in diameter and is no less than 500 metres deep.

In historical times, and even very recently, only two cases of the fall of large meteorites have been recorded, namely the fall of the Siberian or Tunguska object and the fall of the Sikhote-Alin meteorite.

The Tunguska Object

Unfortunately, in this as in other cases, there were no scientifically trained observers of the phenomenon. Unfortunately..., perhaps fortunately for the supposed observers. A witness, an Evenki herdsman, was thrown high into the air by the shock wave, and then was thrown to the ground, as by a bomb. It was said of the poor man that the blow and the fright left him speechless, and when L. A. Kulik, an investigator of the fall, found the man, this the most valuable witness was unable to give his evidence. The fall occurred on 30 June 1908 in a remote, swampy taiga, near the river Podkamennaya Tunguska, hundreds of kilometres from the nearest railway. It did not attract the attention of the tzarist government, and the scientific investigation did not begin until after the Great October Revolution.

The immensity of the phenomenon and the fact that it occurred in our time allowed many interesting facts connected with its occurrence to be clarified.

The bright fireball was noticed in a number of populated areas of Central Siberia in the clear weather. About 7 o'clock in the morning it penetrated the upper atmosphere and fell toward the Earth in a north-easterly direction. In full sunlight it attracted the attention of the passengers of a train, who were looking out of the windows of the carriages as they were travelling along the recently completed Trans-Siberian railway. It would be interesting if we too could someday be witnesses of such a flight!

The blue-white fireball flew with a velocity of several tens of kilometres per second, constantly approaching the horizon and leaving a wide track across the sky. It lost little of its speed from air resistance and, having travelled half a thousand kilometres across the sky, ended its journey near the river Khushma, a tributary of the Podkamennaya Tunguska which empties its waters into the great Enisey.

People saw a fountain of the products of the resultant explosion as a huge column of smoke 450 kilometres away. In order to be seen from there, it must have risen to a height of not less than 20 kilometres.

In a small settlement 60 kilometres from the fall (the closest populated point) the shock wave wreaked enormous damage. It seemed that half the sky was filled with flame in an instant. This flare, in spite of the bright sunny day, was seen in settlements along the Tunguska and even along the Lena, in gold mining camps hundreds of kilometres from the site of the fall.

A shock wave always turns into a sound wave; so it was in this case. In these settlements, windows and dishes on the shelves rattled from the shock wave, and a weak sound was heard 700 kilometres away. Farther away, where the population paid no attention, it was noticed by instruments which registered the air pressure. These instruments – barographs – noted the air wave in St. Petersburg, Copenhagen, Potsdam, and even Washington. From these instruments' record the moment the air wave arrived can be determined. The wave was traced back to show how it propagated from the Podkamennaya Tunguska east and west, farther and farther. Travelling around the world and becoming ever weaker, it still continued to travel, and 30 hours later was registered for a second time in Potsdam.

A shock wave of this sort has only been registered one other time, when the volcano Krakatoa erupted in 1883, completely destroying the island which lay at its base.

Just imagine the might of the earthquake and sound created by the fall of a giant object. Anyway, it weighed thousands of times more than the largest bomb and was hundreds of times faster than a bomb released from an airplane. The shock wave of the earthquake caused by the fall (seismic waves) immediately travelled through the Earth's crust becoming attenuated with distance from the centre of the explosion. They were registered by automatic earthquake recording devices (seismographs) in Irkutsk, Tashkent, Tbilisi and even in Germany. True, in Germany, the ground moved only a small fraction of a millimetre, but accurate instruments such as the seismographs can register even this movement.

But what occurred at the location of the fall?

The small mountains and thick forest around the location weakened the effects of the shock wave, but still the tents and huts of the Evenki herdsmen were torn from the ground as if by a whirlwind, and the residents were knocked down and injured. The tents were 30 kilometres from the fall!

Leonid Kulik, an Associate of the Academy of Sciences, was the pioneer in the study of the fall at the Tunguska. Through barely passable taiga, thick forest, swamp, in the midst of clouds of constant, exhausting midges, initially with scanty supplied (since it would have been impossible to take a great deal on such a difficult journey), he searched over a hundred square kilometres. Launching himself into the backwoods indicated by the local residents, he discovered that the explosion encompassed an area about 25 kilometres across. Tree-branches in this area were burned off their trunks, which were themselves scorched 1-2 centimetres below the surface. People said that in addition to the burnt trees, deer had been found roasted alive there. Although these stories are anecdotical, there can be no doubt that the heated air destroyed everything living in the area at the time. The shock wave broke off trees within scores of kilometres around, the tops lying pointed away from the direction of the explosion. An aerial survey made in 1937-1938 showed that there were actually two centres of destruction of the forest.

During the years 1927-1930 Kulik discovered that the peat covering the swampy soil there was piled into mounds several metres high by the air pressure and in places had been broken into pieces and dislodged. Small pieces of powdered rock were found in the clay, carried there by the explosion. Not far away a destroyed Tungus store was found. Also, ten more craters ranging from 10 to 50 metres across were found, as well as pieces of melted quartz with traces of nickel iron, but they were not associated with this fall.

Apparently, the meteorite broke up and fell in several pieces over a rather wide area. These pieces, 20 to 30 years since the fall, could have penetrated deep into the swamp and become covered by mud or overgrown by vegetation. The craters had nothing to do with the meteorite.

The fact is that the Tunguska object fell into a permafrost region, where the frozen soil at a certain depth never thaws. The permafrost layer does not let water pass through, and subsoil water freezes at a small depth, raising the upper layers of the soil into hillocks. The small craters were formed of lines of these hills.

The difficulty of travel and the uncertainty of success delayed progress. It was decided first to clarify the situation with aerial photographs and then make the large expenditures required to search for meteorite particles. Just before the war the photographs were studied, but the Second World War interrupted further scientific explorations, just as they disrupted almost every aspect of peaceful scientific life in our country. Kulik volunteered for the Red Army and died without seeing the realization of his dream – the location of the Tunguska object.

What is the answer to the enigma of the Tunguska object? A science fiction story portrayed the Tunguska object as an interplanetary ship from Venus that exploded without a trace when it crashed, due to stores of atomic energy on board. This imaginary story was taken seriously by some people, but the only truth in it is that there was indeed an explosion. But the explosion did not require either the residents of Venus or atomic energy, nor even astronauts from Mars, to whom some science fiction writers have also attributed an attempt to land in the wilderness of the taiga.

As was demonstrated by the calculations of K. P. Stanyukovich and V. V. Fedynsky, the largest of the meteorites, such as the Tunguska and Arizona meteorites, reach the surface of the Earth without losing their speed. Thus, even at a velocity of 4-5 kilometres per second a solid body acts at the moment of impact as if it were a highly compressed gas. The crystalline lattice of the meteorite disintegrates instantaneously and turns into gas, which then expands violently.

In this manner a very real explosion is produced, resulting in unbelievable devastation. The meteorite itself disappears, being converted into gas and distributed through the atmosphere. The particles that actually fall can only be companions of the meteorite, which have separated from it before falling to the ground, and, owing to their small mass, have moved through the atmosphere more slowly.

Finally, in 1957 microscopic particles of meteorite iron were found in the soil, although they are also encountered in other areas of the Earth.

So, there is no longer any mystery of the Tunguska object: it exploded before hitting the ground and was transformed into small particles, dust, and even gas.

Fesenkov considered that the fall was not due to a meteorite, but to the core of a small comet. This does not change the essence of the case, however. The meteorite (or the ice-stone core of a comet) did explode, and therefore its remains will not be found.

It has now been established in general that when meteorites fall slowly impact craters are formed, and when they fall quickly and explode in the air, explosion craters are produced with the result that a meteorite may even vaporize completely.

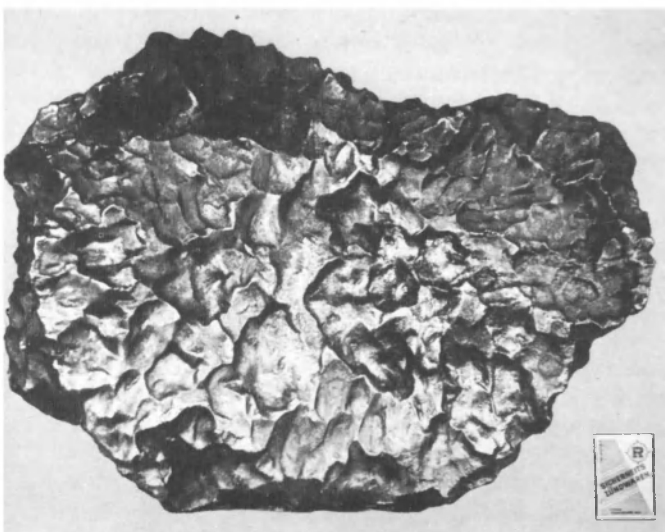
Sikhote-Alin Meteorite

In the Primorski Krai (Maritime Territory of the USSR), which lies on the shores of the Pacific, rises the picturesque Sikhote-Alin mountain range, described by the Russian writer and traveller Vladimir Arsen'yev (1872-1930).

In April 1947 from Vladivostok a group of miners and military engineers left for the fall site wending their way through the forested hills, making a long trail in the snow reminiscent of a snake. After them Academician Fesenkov and his wife E. V. Pyaskovskaya, who was a geophysicist, came striking their way through snow drifts. This unusual expedition was not prospecting for minerals, it had come in search for some of the iron meteorites that had fallen in a shower in the taiga cratering the ground there. The engineers in the expedition were there to blast any of the stones covering the craters and remove the remains of the heavenly guests that might lie beneath them.

The meteorite shower had fallen on 12 February 1947 and was the largest and best studied. The flight of a bright fireball with a varicoloured tail was seen by many residents of the Primorski Krai within about 300 kilometres. The fireball disappeared behind the mountain range and after a few minutes loud explosions came, followed by thunder that reverberated through the snow-covered pines for a long time. The smoke trail that remained after the flight of the fireball was seen in the sky all day.

The unusual phenomena – door flying open, plaster falling, animals panicking – made a strong impression on those who did not know what was going on and who had found themselves in the path of the flight of the meteorite. On the third day two aircraft pilots noticed damaged areas of forest and fresh craters in the rocky terrain from an altitude of 700 metres. Thus aviation saved scientist from a long and difficult search for the meteorite in the taiga.



A 125-kilogramme iron meteorite from the Sikhote-Alin meteorite shower.

Even so, the geologist F. K. Shipulin had to make his way for 100 kilometres through the snow in search of the fall site. Somewhat later, the expedition sponsored by the USSR Academy of Sciences arrived with Academicians Fesenkov and Krinov, who settled in a hut constructed by the engineers, undisturbed that the engineers had not finished erecting the roof. The engineers and the scientists worked harmoniously together in the hills, in spite of the difficult conditions.

The meteorite broke into thousands of pieces at the very last moment and fell uprooting, breaking, splitting and damaging the trees in an area 5 × 20 kilometres about the impact site. The larger pieces fell at the head, forward portion of the devastated strip. All in all, 106 craters were found ranging from 0.6 to 28 metres in diameter. The deepest crater was 6 metres deep. Several thousand fragments weighing from 1,745 kilogrammes down to fractions of a gramme were collected inside these craters, from their edges, in the mud thrown out of the craters or piled up within them, in the soil and even in thick leaves of surrounding plants. Altogether, about 23 tonnes of meteoric iron was collected, but obviously even larger amounts remained undiscovered or had vaporized on impact since the mass of the entire shower was estimated to have been 100 tonnes (and before it penetrated into the atmosphere it could have been 1,000 tonnes).

The shape and structure of the meteorite fragments as well as the destruction they caused have given us a great deal of evidence that can be used to increase our understanding of how meteorites fragment upon their entry into the atmosphere. The meteorite entered the Earth's atmosphere at about 20 kilometres per second, made it down to about 5 kilometres above the ground before being broken up into pieces. There was no explosion when the meteorite hit the surface, as is normally the case when a meteorite has a large mass and is moving too speedily for the atmosphere to reduce its cosmic velocity to that of an ordinary falling body.

Academician Fesenkov managed to calculate the orbit of the Sikhote-Alin meteorite from observations of its flight through the atmosphere. It appeared to be similar to the orbit of either a short-period comet or one of the small asteroids whose orbits intersect those of the inner planets (e. g. Icarus). Fesenkov thought at first the Sikhote-Alin meteorite was an asteroid, one of the many that have never been observed because they are too small. Later he changed his mind and came to think that it was the nucleus of a small comet.

Bombardments from the Heavens

In order to complete our list of the meteorite craters known, we will include a few doubtful cases. For example, the great Ashanti depression, which is 10.4 kilometres in diameter and lies in the Gold Coast of West Africa is suspected to have been caused by a meteorite, as is that of the yet larger (19 kilometres) Ngorongoro crater in East Africa, but neither of these has been adequately studied. If we discount these then all the craters now known to be of meteoric origin may be summed up in the following table:

Location	Number of craters	Size of largest, metres	Date of discovery
Arizona (USA)	1	1,200	1891
Henbury (Australia)	10	200 × 110	1931
Box Hole (Australia)	1	175	1937
Odessa (USA)	1	170	1921
Wabar (Arabia)	2	100	1932
Saaremaa (USSR)	6	100	1927
Campo del Cielo (Argentina)	many	75	?
Dalgaranga (Australia)	1	70	1923
Brenham (USA)	1	17	1933
Sikhote-Alin (USSR)	more than 100	28	1947
Chubb (Canada)	1	3,500	1946
Brent (Canada)	1	3,200	about 1960

Most of the craters were found in deserts where there is little rain, which quickly destroys the contours of these craters raised like an embankments. It is difficult for small craters to remain intact near cities and it is unlikely that they will be found there, since cities are always located where there is some rainfall and the population always attempts to make the most use of every square inch of land.

In each of the cases given here the meteoric origin of the craters was questioned until pieces of a meteorite have been found, invariably iron. The presence of kieselguhr and melted rock and, where there is sand, the presence of glass (melted and then solidified) filaments and splatters are signs of the meteoric origin of craters.

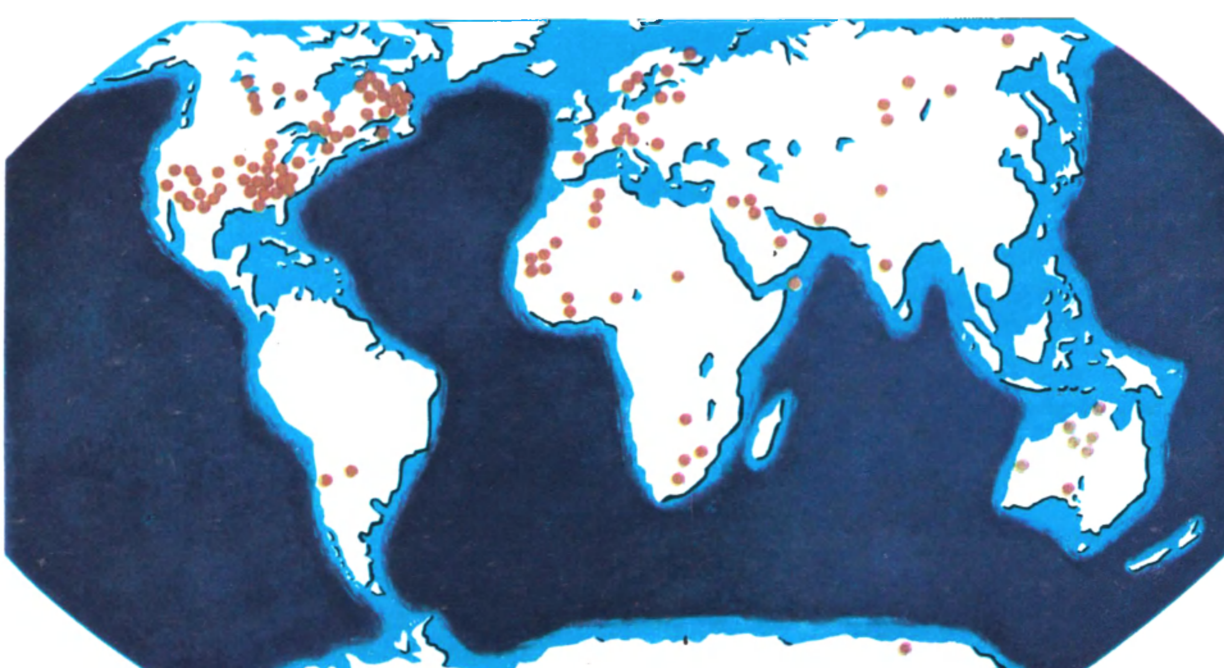
Taking account of the weathering of meteoric craters, we conclude that meteorites capable of forming craters greater than ten metres in diameter collide with the Earth no more than once a century. Apparently, no meteorite ten times larger than the Arizona meteorite has fallen on the Earth, otherwise a crater would have been formed which would have remained intact until today, even if the meteorite had fallen hundreds of thousands of years ago. This confirms our earlier conclusion that there is an extremely small probability of the Earth colliding with the nucleus of a comet and shows that such a collision would be of no danger, not only to the Earth as a whole, but even to individual regions of it. You may sleep peacefully after seeing a shining comet in the heavens approaching the Earth...

The discovery of craters of meteoric origin on the Earth has made many people believe that the lunar ringed mountains – those "pock marks" on the face of the Moon – were formed by meteorites.

When craters were found on small moons of the planets, where volcanoes were impossible, no doubt remained that the craters were due to impacts.

There is one lunar feature that, although it may be the only one of its kind, completely excludes the possibility of participation of such a surface shaper as a meteorite. It is the Wargentin, a flat mountain absolutely circular in plan, without any depressions in the centre but with slopes just like those of the ring mountains. It could only have been formed by lava rising up from the Moon's interior to solidify at the level of the circular rim.

Geography of meteorite craters.



To summarize, the lunar mountains are undoubtedly unique manifestations both of volcanic activity and of meteorite falls in earlier times. And both are responsible for the Moon having a "pock-marked face".

Zodiacal Light and the Gas Tail of the Earth

Still another natural phenomenon is closely connected with the small interplanetary particles we call meteors. Inhabitants of tropical countries are familiar with an occurrence that will only poorly be appreciated from the following description.

On spring evenings, after sun has set and the twilight has faded and on autumn mornings before dawn, a sort of silver cone of light rises from beyond the horizon. In the southern portions of the USSR it is brighter than the Milky Way though it is poorly visible in middle latitudes, because the axis of the glow stretches along the ecliptic (along the zodiacal constellations) and in the middle latitudes the ecliptic only grazes the horizon. Atmospheric absorption and dust interfere with our view of the lower, brighter portion of the zodiacal cone. Sometimes it can be seen stretching from east to west, forming a gigantic arc of light in the southern portion of the sky. At the apex of the arc, i. e. at the point opposite the Sun, the light is brighter and a point several degrees across is seen. It is called *counterglow*.

Since the spectrum of zodiacal light was measured quite recently, it is now well established that it is reflection of the sunlight from countless meteoric particles which fill the space between the Sun and the Earth. The spectrum of zodiacal light is indeed a copy of the solar spectrum, i. e. it is the sunshine reflected by rather large particles. The sunlight could be diffused by tiny dust particles and gas molecules, but then zodiacal light would be bluish, as is the sky colour, which really is created by the dispersion of the sunlight by tiny particles and gas molecules in the atmosphere. But zodiacal light is white and the blue part of its spectrum is not brighter than the Sun's.

No one has studied the zodiacal light in as great detail as Fesenkov, who became interested in it when a student. Having performed many measurements of the brightness of zodiacal light and as a result of complex and detailed analysis of these observations, Fesenkov came to the following conclusions.

Most of the particles that give rise to zodiacal light are very small. The amount of light they reflect can be used to determine their total surface area. A small mass, divided into many tiny particles, has a surface many times greater than the surface of the same mass collected into a sphere. If this is not completely clear, try, for example, to compare how much paint would be required to paint a boulder, and how much would be required in order to paint the hundred or thousand pebbles into which it could be broken.

The meteoric particles are farther away from the Sun than the orbit of the Earth, and those which are directly opposite the Sun reflect the light of the Sun with their entire surface, like a full Moon. In this direction, then, the light is reflected more brightly to form the spot called the counterglow. The meteoric particles located to the side of the Earth and the Sun reflect less light and in total, as can be shown by calculation, should have brightnesses corresponding to the apparent brightness of zodiacal light.

The meteoric particles that form zodiacal light move around the Sun, apparently in random orbits, many of which are sharply inclined to the ecliptic. The majority of the particles move near the plane of the Earth's orbit, so that they form near the Sun a strongly flattened swarm elongated along the plane of the solar system in a lens-like manner, the Sun being in the centre of the "lens". The particles are formed when the asteroids disintegrate and collide with large meteorites; they are being formed even now. They continually fall into the Sun, so that every 100,000 years the Sun receives the matter contained within a sphere with a radius of one astronomical unit. The material in this sphere could make up an asteroid 10 km in diameter.

By studying the distribution of brightness of the night sky in Alma-Ata, taking into consideration zodiacal light, Fesenkov and his associates in 1949 supported the suggestion of Astapovich that the Earth's atmosphere, and the Earth itself, has a shining gaseous tail. This tail, like that of a comet, always points away from the Sun, and spreads over the plane of the ecliptic. Here the Earth's atmosphere, apparently, stretches out into space in the form of a cone with an apex angle of about 10°. The density of the air

The Pleiades, star cluster NGC 1434, Messier 45.

in the tail falls off rather slowly, halving approximately every 4.7 Earth's radii. As in the case of comets, this may be caused by the pressure of solar radiation, although in general the tail lies within the Earth's shadow. This all, however, needs more verification because the conclusions are based on observations of phenomena which are difficult to observe and no possible reason for this behaviour of the air molecules, consisting in the main of nitrogen which should be largely unaffected by the pressure of sunlight, especially in the shadow of the Earth, has been put forward.

Light and Dark Nebulae

In the solar system meteorites are responsible for the pleasures of observing fireballs, shooting stars, for meteorite craters on the Earth, Moon, Mercury, and Mars, for zodiacal light, and even for the rings of Saturn, Uranus, and Jupiter. But this is not the whole of the story and we will again encounter meteorites beyond the solar system.

On a clear, frosty night, under the three bright stars that form the "belt" of the mythological hunter Orion (or the three Magi as the ancient Slavs called them) you can see through the binoculars a twinkling spot of light, silver-green mist. It envelopes the centre star forming Orion's sword-belt.

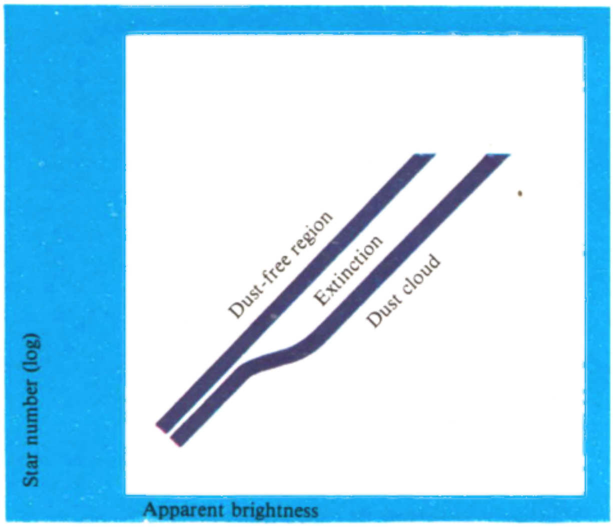
Many such diffuse light spots were discovered by the telescope during the three centuries since its invention, but a whole world of these nebulae opened up before us with the employment of photography for astronomical purposes.

Accumulating the light for minutes, sometimes hours on end, the photographic plate reveals the presence of these weakly shining points, nebulae, some of which cannot be seen even with the aid of a telescope.

Near the Milky Way, and especially within it, we find a multitude of shining, ragged nebulae, large and small, named diffuse nebulae because of their lack of form. All are entered in the catalogue. Their spectra tell us that some consist of rarefied gas, others of something else, which reflects the light of the stars.

In 1921 the American astronomer Hubble discovered that if the spectrum of a diffuse nebula is continuous and has dark lines, it is always an exact copy

The Pleiades (contrast), photographed on 3/4 January 1962.



The variation of the number of stars with respect to their apparent brightness in dust-free region and dark cloud.

of the spectrum of a nearby star. In the majority of cases the star is visible against the nebula, and this is almost certainly not a chance occurrence, not a projection of the star on the nebula, but evidence that the star is either near the nebula, or perhaps within it.

One of the best examples of this is the group of weak nebulae which surround the brightest stars in the star cluster Pleiades. The spectra of the nebula and the changes in their brightness with increasing distance from the stars indicate that these nebulae shine with light reflected from the stars.

But what could be collected into such chaotic, formless clouds and reflect the light of the stars, other than cosmic dust? In some cases the spectra of nebulae consist of a combination of a sharp-line spectrum and the usual star spectrum, thus indicating a mixture of thin gases and meteorite dust. Almost always the bright nebulae are accompanied by dark spots. In the mass of shining material are visible black spots and bands like ink blots on thin white gauze.

Originally, these dark spots did not attract the attention of astronomers, but later Barnard in the USA found many of them on the shining background of nebulae. One such spot, perhaps the most remarkable, is located in the Milky Way in the constellation Orion, directly next to the famous Orion Nebula.

Some people, finding it to resemble a horse head, named it also the Horsehead Nebula.

Notice how in the upper part of the photograph the entire field is lit by gentle radiance from the light fog and the great multitude of weak and apparently distant stars. Below this is yawning blackness, against which only the brightest stars can be seen. In the southern skies, in the Milky Way, there is a "Great Hole", a dark spot called by mariners the "Coalsack".

For a long time these dark spots were thought to be "holes" in the mass of the stars that form the Milky Way, empty places in which, for some reason, no stars existed. This explanation once helped a resourceful astronomer to satisfy a female V.I.P. visitor to the observatory. He showed her Mars with its white polar caps, showed her the long shadows of the lunar mountains, indicated the amazing rings of Saturn, but to no avail. The visitor was not satisfied and insisted that she be shown the "endless reaches of Outer Space" through the telescope. To please the rich woman, the astronomer turned the telescope towards one of the dark spots in the Milky Way with its few, barely shining stars. "Look between those stars into the blackness," he told her, "and you will see the endlessness of Outer Space you want so much to see."

That was decades ago. Now such a request would be harder for the astronomer to satisfy. Since that time, the seemingly empty spaces have been "filled" with material... First it was guessed that these were not holes, rather dark curtains covering not only "endless space", but also relatively close stars in our own stellar system.

The curtain is made up of colossal clouds of dust, blocking and diffusing the light of the stars located both within the clouds and behind them. The larger and denser the cloud, the more light it absorbs, and the closer it is to us, the fewer stars we see in that direction. Starlight is greatly weakened by passes through the cloud, so that the stars seem weaker than they really are, weaker than if we saw them through empty space. The absorption makes the weaker stars completely invisible. But, in general, the weaker a star, the farther it is from us, and the farther away, the more stars there are. Therefore, the weakest and most numerous (farthest) stars in the dark spots in the Milky Way cannot be seen, and these spots are now called dark nebulae.

The "curtain" theory precipitated the development of methods of checking it, and these confirmed the existence of the many dark nebulae that absorb light and consist of cosmic dust. Ambartsumyan and Gordeladze proved that these clouds of cosmic dust seem to be dark when there are no bright stars near them and light when they are illuminated by a bright star, which happens to be in the area.

The basic method of studying dark nebulae (their dimensions, distance from Earth and absorptive capacity) is, as was first shown by the German scientist Wolf, to count the number of stars of a given brightness within and around the nebula.

A dark dust cloud may be extremely thick, often measuring dozens and sometimes even hundreds of light years across greatly exceeding the average distance between stars. Thus, thousands of stars are lost, like birds in a fog, in each nebula.

The absorptive capacity of a dark nebula is determined by the magnitude of the largest light absorption effect it produces. If, for example, in a nebula the number of stars of the 15th magnitude is equal to the number of stars of the 14th magnitude around the nebula (per unit visible area), then the absorptive index of the nebula is equal to one unit ($15 - 14 = 1$) of stellar magnitude. In other words, the nebula weakens the light of stars located behind it by a factor of 2.5.

The area of the sky visible to an observer is determined by the cone receding into infinity with the point at the eye of the observer. The farther from us, the greater the volume it encompasses and the greater the number of stars that should be contained in it. This explains why the number of stars with a given brightness per unit *visible area* of the sky increases with decreasing brightness, i. e. with increasing distance. At a distance corresponding to the average distance to 12th magnitude stars a dark nebula begins, which reduces the number of weaker stars visible. But this method of investigating nebulae is very inaccurate. In recent times, dark nebulae have been studied by more complex but more accurate methods that take into account the variations in the true brightnesses of the stars and a number of other considerations.

The nearest dark nebulae are 300 light-years away from us. As was shown by the calculations of Academician Fesenkov and Professor Parenago, even though they have enormously great dimensions, dark nebulae (on average 3 parsecs across) do not have such great masses; on average a nebula's mass is three times

The Horsehead nebula in Orion. MGC 2024.

that of the Sun, if it consists of small dust particles.

If many small particles can reflect more light than one body of the same mass (remember the rings of Saturn), then they also are capable of absorbing more light. Just think of the mass of a cloud of smoke in which a smoker practically hides himself. If the smoke were gathered into a solid sphere, we would hardly be able to see it, it would be so small that the smoker could never hide in it. The great reflective power of small particles is also the idea behind smoke screens used for camouflage in wartime.

Calculations show that the greatest absorption (or more accurately, diffusion) is attained with particles around one ten-thousandth of a millimetre in diameter, or a little over the wavelength of green light. A column one square centimetre in cross section contain-

ing only 0.1 milligramme of these particles is practically opaque. It would weaken light passing through it by 9 stellar magnitudes, or 4,000 times!

With particles less than one ten-thousandth of a millimetre across, the absorption becomes selective, as with the gas molecules in the air, i. e. the absorption is the greater the shorter the wavelength of the light. This is the reason for the blue colour of the sky and the red colour of the setting sun. The blue rays are diffused the most, whilst the red light passes through the atmosphere very easily. Hence the predominance of red in the light that must pass through the atmosphere, thickest in the direction of the horizon.

If the cosmic dust (for example, metallic iron particles) in dark nebulae is very fine, it produces the same reddening effect on the light of stars located behind it, and the degree of the effect allows the size of the particles to be estimated. This selective absorption of light is discovered by comparing the colour of stars of the same temperature, lying behind the nebula and beyond it. The true colour of a star being dependent on its temperature, stars whose light has been selectively absorbed show an excess of red. Similarly, our Sun, seen through a cloud of dust or fog, would seem redder than it actually is.

The excess of red (often called simply colour excess) and the colour of a star are most accurately measured with the help of the photoelectric photometer. A star's brightness is measured using a photometer directly and also through a yellow or red glasses (light filters), which pass the yellow and red, or just red, rays the best. This gives the share of red in the total spectrum of the star. Some dark nebulae cause colour excesses greater than a unit of stellar magnitude with the result that a white star with a temperature of 15,000 K, say, appears as absolutely red when seen through such a nebula, just like a star of only 3,000 K.

Colour excess and brightness studies of stars in the direction of a dark nebula provide the best way of determining the distance to, size of, and particle size of the nebula. In addition to the fine meteoric dust, dark nebulae should also contain large particles, which would absorb light by simply blocking it. There may be many of these, they may be as large as those that produce shooting stars on Earth. Some may even be as large as those that, having arrived in the solar system, fall to Earth as heavy pieces.

Interstellar Wastes

In Galileo's time, there were many who believed in the purity of the skies – the space over our heads – that, in their opinion, could not have been soiled by the foot of man, who had not been there. This belief caused them vehemently to deny the existence of spots on the surface of the Sun, the emblem of heavenly purity. Science denied these unfounded fantasies, and now not only has the existence of sunspots been confirmed, but science has also proved the presence of fragments and splinters of matter throughout interplanetary space, and even dust nebulae between stars.

But the rest of the space between stars should be clean, or so astrophysicists thought not long ago. This idea was not the result of unfounded beliefs, it was based on evidence. For example, some stars are gathered in tight globular clusters and located farther away from us than the majority of other stars: they show no colour excess. Their light seems only to vary with the square of the distance to them, a dependence that could only exist for an absolutely transparent medium.

The existence of a definite absorption of light, and therefore of an absorbing medium, was first indicated with certainty by the Russian scientist V. Ya. Struve in 1847. He came to this conclusion after studying the distribution of stars in various directions. The conclusion, however, was so audacious and unexpected that other scientists doubted it, and even after that many of them continued to think of interstellar space as completely transparent.

In 1930, Trumpler in the USA, studying dense star clusters, discovered a strange and unlikely fact. The farther a star cluster is away from us, the larger is its linear diameter. It is to be noted that studied were not the globular clusters mentioned above, which can be seen to the side of the Milky Way, but the so-called galactic clusters, which are closer to us than globular clusters and visible in the belt of the Milky Way itself. These include the star clusters Pleiades and Hyades. Their visible position in the Milky Way indicates that they lie near the plane in which the stars of our stellar system are located, thus making around us a ring of stars – the Milky Way. Understandably, before making public his surprising findings, Trumpler tried to think if there could be something behind them that had been overlooked before. He considered once more the method of studying star clusters.

First, the distances to the nearest clusters were determined by methods that were beyond doubt. Knowing these distances, the linear dimensions of these clusters could be derived from their apparent angles. It was found that the linear dimensions and brilliance of the main star clusters of the same type were all the same. It was only natural to decide then that, for clusters of the same type, the apparent brilliance of the main stars only varied with their distance, inversely proportional to its square. Knowing the true brilliance of these stars and comparing it with the apparent brilliances of the stars, which were easy to measure, it was not difficult to calculate how far away the more distant clusters are. From the thus calculated distance of the more remote clusters and their apparent angular diameters their linear dimensions were worked out. And these latter appeared to grow with distance!

What was the matter? If the distances obtained were overestimated, and at that the more, the larger the distances themselves, then the result would be exactly like this: the linear diameters would also be overestimated and the more, the greater the distance. But the apparent brightness can be influenced by distance if the space absorbs light. The farther the cluster, the longer the path through the absorbing medium the radiation must follow, and the greater the reduction in its brightness. Unaware of the absorption, we placed the clusters farther away than they actually were. As Trumpler has finally established, there is some absorbing material in our stellar system; it is concentrated in the plane of the Milky Way, which is in itself a rather thin layer. It was later calculated to be 600 light years in thickness. Light from bodies located within this layer (like our solar system) is strongly absorbed in it (e. g. galactic clusters). On the other hand, light from sources outside this layer (e. g. globular star clusters), undergoes very little absorption, since it goes through the absorbing medium for only a small portion of its journey. It was for this reason, that the absorbing material had been unnoticed when studying globular star clusters.

It has been discovered that the apparent brightness of stars reduces near the plane of the Milky Way every 3,000 light-years by 0.4 units of stellar magnitude, and the brightness of stars in photographs, by as much as 0.7 units of stellar magnitudes; in the plane of the Milky Way itself, the reduction is even greater.

Later investigations have indicated that the interstellar medium is distributed in space very unevenly: in some places it is denser than in others. The absorption weakening the apparent brightness of stars varies with their distance and direction. In some directions, in the plane of the Milky Way, the absorption goes up to several units of stellar magnitudes for each 300 light years.

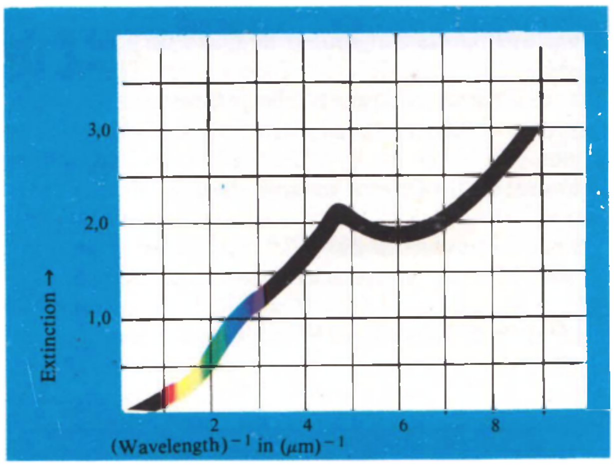
Since our understanding of the structure of our stellar system is based on the number and brightness of the stars, and the latter is affected by the uneven absorption of light in space in different directions, it is necessary to study and consider this absorption constantly. Studying the structure of a stellar system has been made very difficult and complex with the discovery of the light absorption. The stellar system must be examined, so to speak, in pieces, step by step, separately in each direction. Obviously, light is partially absorbed by a continuous dust medium filling the space between stars, and partially by the many distant dark nebulae traversed by the light.

The general absorption of light by interstellar dust is accompanied by selective absorption. The nearer to the plane of the Milky Way the stars lie and the farther they are from us, the redder they are due to selective absorption. In recent years, photography using plates sensitive to red light has revealed whole areas of the sky – whole star clouds in the Milky Way – with a reddish colour due to the selective absorption. Ordinary film, insensitive to red light, does not show these distant clouds of weak stars very well, but on red-sensitive plates they come out more clearly and many more stars can be seen.

The absorbing material in the Milky Way is not a thin, sharply bounded layer, its density rapidly increases towards the plane of the Milky Way, just as that of the Earth's atmosphere thins out with altitude.

The centre of our stellar system contains the thickest clouds of stars and is oriented in the direction of the constellation Archer in the Milky Way. It is hidden from us by the absorbing material. If it were not for this, the star clouds in this direction would shine with almost blinding brilliance.

The dark "fork" in the Milky Way, which divides it into two branches beginning in the constellation Swan and reuniting in the southern hemisphere of the sky, is also formed by dark nebulae and is not a giant hole in



The dependence of light extinction in interstellar dust on wavelength, enabling the dust particle size to be determined.

the Galaxy. In the direction of this fork, as in the other directions where absorption is especially great, we do not see any distant star formations, lying beyond the Milky Way. We see there neither globular star clusters, nor any of the so-called spiral galaxies, which are covered in the sequel volume. They are hidden from our view by the dark interstellar material. However, in places there are a few "windows" in the dust mass of the Milky Way, like polynyas in the Arctic ice mass or windows of open water in the green of an old pond. Through the windows can be seen stellar systems lying beyond the limits of our Galaxy. The author has also discovered "visibility corridors" in which distant stars of our own Galaxy can be seen, although other, yet farther, stellar systems cannot be seen. In the directions perpendicular to the plane of the Milky Way, for example in the constellation Veronica's Hair, space is almost completely transparent. There we can easily see the most distant corners of the Universe, those lying far, far beyond the confines of our Galaxy.

Light is known to be electromagnetic oscillations perpendicular to the direction of its propagation. Normally, the plane of the oscillations changes continuously and randomly. Light is, however, made more orderly by being passed through certain crystals or reflected from certain surfaces. In the process, one plane of oscillation becomes predominant. This phenomenon is called *polarization*, and the light is called *polarized*. Using special instruments it is possible to determine if light is polarized or not. Starlight, which has passed through a dust medium, is polarized the more, the farther the star is away from us.

The polarization of starlight by interstellar dust indicates that the particles are not round, but elongated, and that they are not distributed randomly in space, but are near the plane of the Milky Way. What makes them do this? Apparently, this is the result of the presence of a magnetic field in outer space. Thus, the study of the influence of interstellar dust on starlight led to the discovery of magnetic forces in interstellar space.

Light absorbing dust is found not only in our stellar system. In stellar systems similar to the Galaxy, we find some light absorbing material concentrated again in the planes of symmetry. This is best seen on photographs of those galaxies which are turned edge-on to us and look like lenses or spindles. In many such photographs a sharp, thin, dark band is visible, which seems to divide the lens or spindle in half. There is no doubt that this dark band is simply a collection of dark nebulae, consisting of cosmic dust, densest near the equatorial plane of the systems. They absorb the light of the myriad of stars within them just as the nebulae do in our stellar system, and the blackness which they create speaks clearly of the greatness of their absorption.

Thus, the investigations of the past decades have discovered the universality of meteoric material, its distribution throughout the Universe, beginning with the solar system and ending with distant stellar systems.

Now, having seen that dust is exceedingly abundant in the Universe, and that it is one of its basic materials, we should be prepared for the fact that at least some worlds are formed from this material, although in interstellar space there also exists an enormous quantity of thin gases, which can also be material that forms other cosmic objects.

When the astronomer shows you a dark spot in the shining Milky Way through his telescope, do not think that it is "endless space", without limit and consisting of "nothing". Rather consider it a cradle in which a new world is being formed. Possibly, in the future it is destined to become a planet, like the one which has given us life and from which we seek to unravel secrets of life that are in store.

Contents (Part 2)

PART 2

The World of Gas

In the part of the Universe of which we have knowledge, gas is the predominant form of matter. In space, thin gas forms clouds of enormous dimensions, called nebulae, and, when concentrated into enormous hot masses, it forms spheres similar to our Sun. Thin gas fills the space between the stars.

The structure of matter in the form of gas is simpler than in solid bodies. In space, gases consist of atoms and very simple molecules; there are no minerals, rocks, or organic matter.

The study of celestial gaseous bodies greatly increase our understanding of the structure and behaviour of matter in general. It supplements conclusions reached in physics laboratories since matter can be observed in states and at pressures and temperatures that cannot be created in the laboratory. The study of gigantic celestial bodies, therefore, makes it possible to penetrate the secrets of the tiniest particles of matter, the interiors of atoms. The exploration of the world of gas not only opens up to us the laws governing the development of the material world but is also of practical significance since only with a thorough knowledge of the laws of nature can man become its master.

1. The Sun – The Closest of the Stars

Introduction

The Sun: life-giving, bright, beautiful, radiant! How many such epithets can be applied! You are the source of life on Earth and the head of the planetary family, you are the closest star. The Sun can be discussed from the point of view of the meteorologist, the radio physicist, the physician, the botanist, the chemist, and the poet – not to mention the astronomer. Many phenomena have been discovered on the Sun, and the descriptions of them often resemble those given at a meteorological observatory when changes of weather or complex movements of atmospheric masses are recorded. Perhaps in the future, and we hope it will be the very near future, scientists will give a theoretically complete explanation of these facts and fit them into a well-organized picture of the physical nature of the Sun.

We observe turbulent changes on the Sun but their reasons are not entirely clear, although the general structure of the Sun has already been known to us for quite some time, and in recent years the theory of solar phenomena has advanced considerably.

Since the Sun is closer to us than the other stars, it can be investigated in considerable detail. The results of such study will be of help in clarifying the nature of other distant suns, visible only as bright dots, even in the strongest telescopes. Let us take a closer look at the Sun as a representative of the world of countless stars.

An infernally hot gaseous sphere radiating fluxes of heat and light, the only source of them in the solar system, such is the closest of the stars.

We know of the Sun's energy from that part of it which comes to the Earth from a distance of 1,500,000 kilometres. Even making allowances for atmospheric absorption, a square centimetre of the Earth's surface perpendicular to the Sun's rays in one minute receives enough energy to raise the temperature of 1.94 grammes of water by 1 degree. This value varies somewhat with the Earth-to-Sun distance, and is probably influenced by phenomena occurring on our luminary. Before scientists suspected that solar radiation might vary, this value was given the name "solar constant". Geophysicists measure it directly almost every day by observing the heating of water in special vessels exposed to the Sun's rays.

Multiplying the solar constant by the value of the surface of a sphere with a radius of 1,500,000 kilometres, we can determine the total radiation of the Sun. This value is $5.43 \cdot 10^{27}$ calories per minute. When converted to mechanical power, it would be equal to $0.37 \cdot 10^{27}$ W. To give a graphic idea of this magnitude, we could say that if we instantaneously encased the Sun in a layer of ice 12 metres thick, it would melt in a minute. If we were to extend a bridge in the form of a column of ice 3 kilometres thick from the Earth to the Sun and concentrate on it all the Sun's radiation, the bridge would melt within one second and it would vaporize within 8 seconds.

Sunlight illuminates the Earth 465,000 times brighter than a full Moon does, and if it were to be replaced it would be necessary to have a light source of 135,000 candelas placed at a distance of 1 metre.

Knowing the distance to the Sun and its apparent angular diameter (1/2°), we can determine its linear dimensions easily. The Sun's diameter is greater than the Earth's by a factor of 109, its surface is greater than the Earth's by a factor of 109^2, or 12,000, and its volume by a factor of 109^3, or 1,300,000. (The solar diameter is 1,390,600 kilometres but this figure is not easily to comprehend.)

By dividing total radiation of the Sun by its surface, we find that one square centimetre gives the same light as 50,000 candelas. The Sun's surface is 10,000 times brighter than molten platinum and is 10 times brighter than the flame of an electric arc. In the total flux of energy emanating from the Sun light amounts to 8 candelas per watt, whereas for ordinary electric incandescent lamps it is no more than 2 candelas per watt. This high luminous efficiency of solar matter is due to its high temperature.

The radiation power from a unit surface of the Sun, obtained by dividing its total power by its surface, is 62,000,000 watts per square metre and, undoubtedly, it has been this way over a period of hundreds or even thousands of millions of years. How stable and powerful the energy sources in the Sun's interior must be! We will discuss this in a later chapter and here we will only note that an insignificant part of the energy generously expended by

the Sun is used by the planets. The fraction received by the Earth is only 1/2,200,000,000th part. Nevertheless, it is colossal. If we assign to it a value of 2 kopecks per kilowatt-hour, we find that every second the Earth receives from the Sun a thousand million ruble worth of energy.

This money is literally scattered to the winds, since the wind itself is a movement of air caused by the unequal heating of different parts of the Earth's atmosphere and the Earth's surface. The greater part of the Sun's energy is turned into the energy of water (the water of rivers and rain). A part of the solar energy is used by plants, and a small part of it is used as fuel in the form of peat, firewood, and coal.

Much has been said about the direct use of solar energy, in addition to that brought to us by the moving force of the wind and water. The solution of this problem involves many difficulties, but clearly engineers concerned with energy problems have so far given it too little of their attention. Mention should be made of the uses of solar energy on artificial satellites of the Earth and Sun, which was made possible with silicon photocells that transform solar energy into electric energy.

The Sphere of Light

The sphere of light, or in the jargon of science, the *photosphere*, is the name given to the visible part of the Sun's surface. In reality, this is not a surface at all, since the Sun's rays reach us from different depths of the mass of gases that make up the Sun. Light coming from the deeper layers is more attenuated, of course, as a result of the incomplete transparency of the upper layers, and beginning at a certain depth cannot escape at all. The solar gases in the colder, outermost layers are especially opaque.

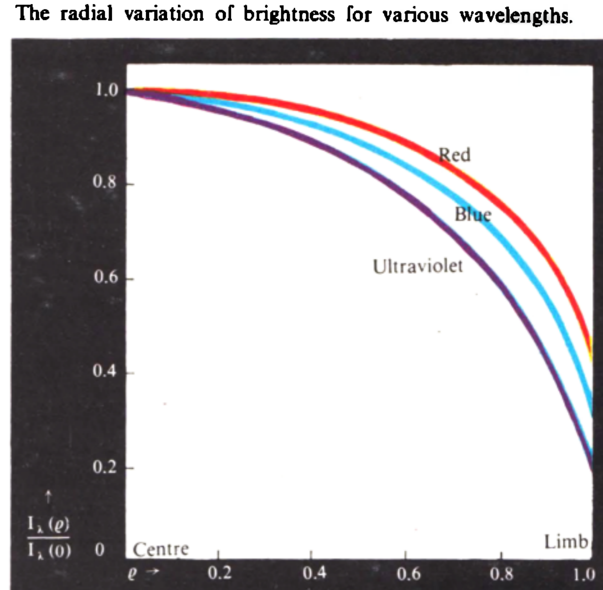
The opacity is caused by negative ions of hydrogen (these are hydrogen atoms that picked up an electron). Therefore, the solar limb appears very sharp and not fuzzy, although there is no sharp jump in the density of solar gases, which could be taken to be the boundary or the surface of the Sun.

Towards the limb the Sun is appreciably darker than at the centre, because of the low translucence of the photospheric layers. The light reaching us from the centre of the Sun comes from deeper and hotter layers. Light from the limb, on the other hand, comes from shallower, cooler, and therefore darker, layers of the photosphere.

To all intents and purposes we cannot see the layers lying deeper than a hundred kilometres

The solar observatory Einsteinturm (Potsdam, GDR).

The radial variation of brightness for various wavelengths.

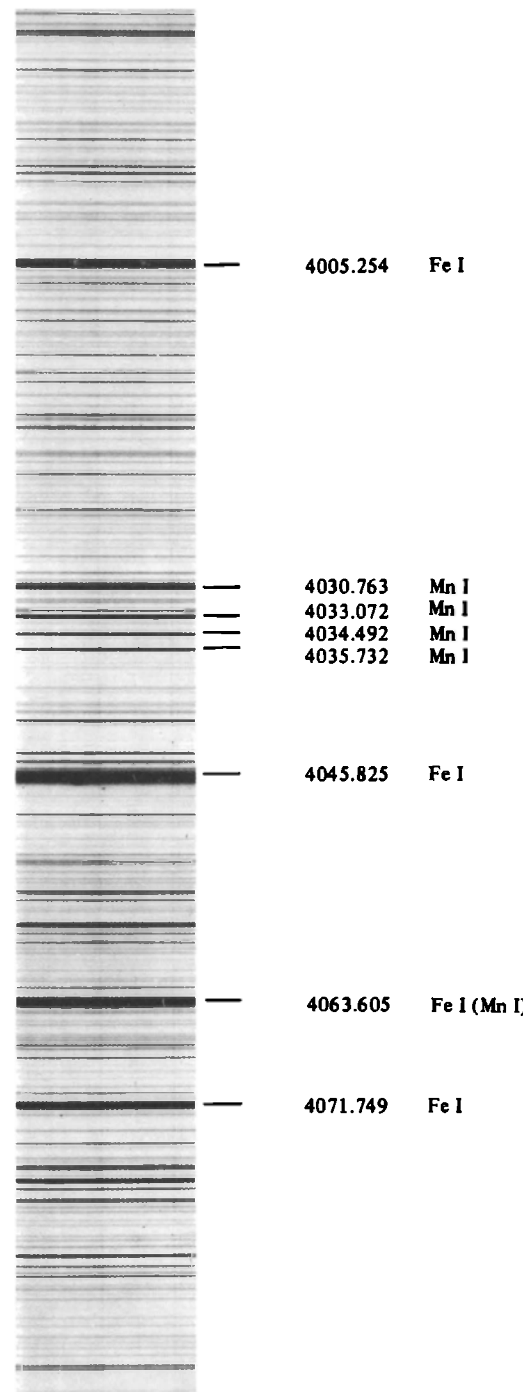


beneath the layer which forms the boundary with the solar atmosphere, although the blindingly bright gases of the photosphere are almost a million times thinner than atmospheric air.

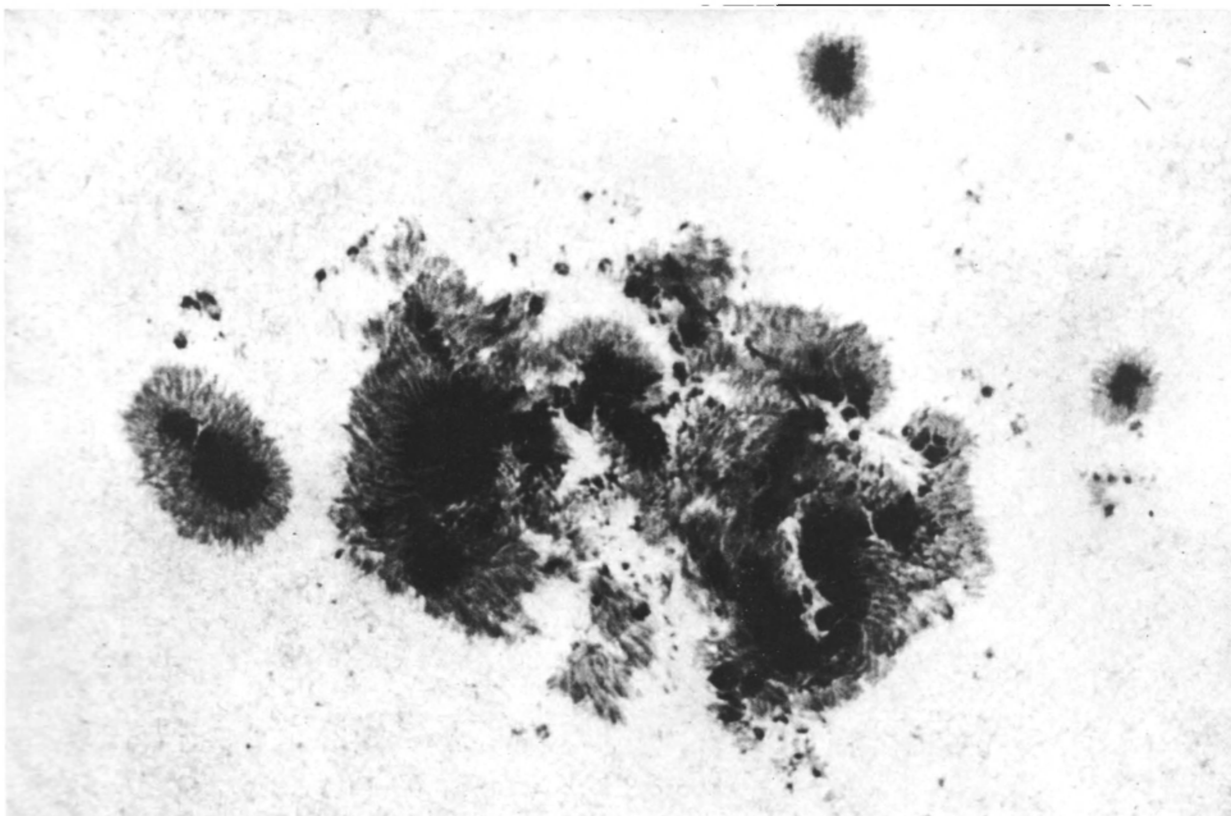
The temperature of the solar "surface" or "photosphere" is an arbitrary concept. The deeper within the Sun, the higher is the temperature of the layers. The temperature of the photosphere, a sort of mean of the temperatures of the outer layers, can be determined by many methods, for example, from the energy distribution over the continuous spectrum that characterizes the radiation of the photosphere. In actuality, the spectrum of the photosphere is a combination of continuous spectra that correspond to the different temperatures of the various outer layers of the Sun. It is this complex spectrum that is the spectrum of the photosphere, but it is quite close to the spectrum of a true black body at a temperature of 6,000 °C, which is assumed to be the temperature of the solar photosphere. It is sometimes said for short that 6,000 °C is the temperature of the Sun, although the temperature and density of the gases of the Sun vary with depth.

Even before these details had been revealed, Professor V. K. Tserasky of Moscow made the following interesting experiment to determine the lower limit of the Sun's temperature. He took a concave mirror 1 metre in diameter to focus the Sun's rays in a solar image with a diameter of 1 centimetre. It can be shown that at the focus the temperature cannot be higher than the Sun's. However, Tserasky melted substances with extremely high melting points. It is not surprising then that at the Sun's surface all matter, including metals, is in a state of hot vapours. Moreover, these layers, which are twice as hot as the flame of an electric arc, are colder than the Sun's interior whence the photosphere derives energy to make up for losses by radiation. The photosphere is a perpetually open door through which the solar energy escapes.

It would be incorrect to think that the solar photosphere is homogeneous. It has a granular structure, and its oblong grains are called granules. These "grains", however, are thousands of kilometres across and consist of hot gases. The granules can be seen clearly in a small telescope. They appear constantly and disappear more rapidly than the



A section of the solar spectrum.



A sun-spot group with granulation.

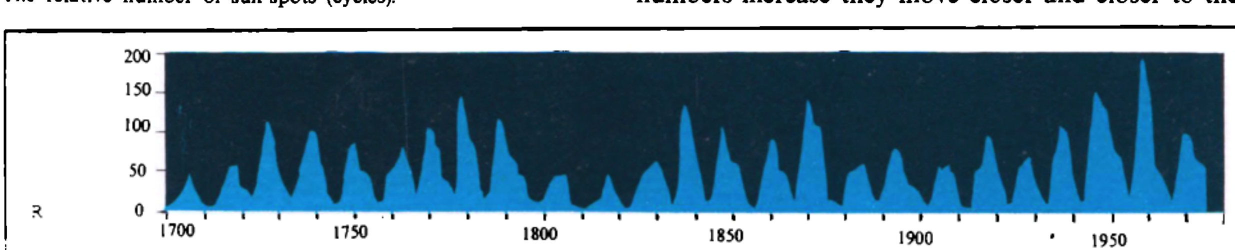
clouds in our atmosphere, existing for only a few minutes. The granulation of the Sun's surface is similar to a boiling rice gruel, so much that when a motion picture is required of the image of the Sun's surface it is this gruel that is photographed. The light coloured granules are rising jets of hot gas. In the dark intervals between them gas is descending and these places are cooler. In this way, convection occurs in the Sun's upper layers and as a result, the heat from the lower layers is carried to the outer layers.

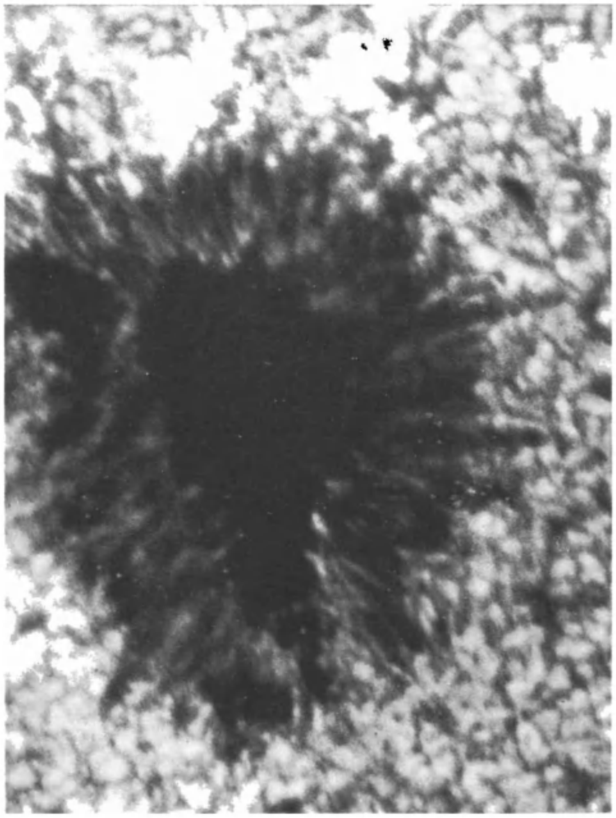
Small bright formations called faculae are often visible, especially at the solar limb where the disk is darker. This indicates that the faculae are cloudlike masses of gas in the upper layers of the photosphere. Their higher apparent brightness at the limb is caused only in part by their higher temperature which, according to measurements made by astronomers at Pulkovo, is only 100 °C higher than that of adjacent lower layers. In the upper layers the faculae are approximately 2,000° hotter than the photosphere. The homogeneity of the photosphere is disrupted to an even greater degree by dark spots that appear on the Sun.

Even the Sun Is Not Without Spots

"Even the Sun is not without spots", contemporaries of their discovery in the 17th century said with vexation. The Sun, it is true, does have spots, but not always. As is well known, the number of sun-spots and the area they occupy change periodically, although not very regularly, with a period of 11 years. In a year of a minimum, not a single spot may appear on the Sun for months. In every cycle, sun-spots appear on both sides of the solar equator at latitudes of about 30°. As their numbers increase they move closer and closer to the

The relative number of sun-spots (cycles).





A sun-spot core (note the granulation).

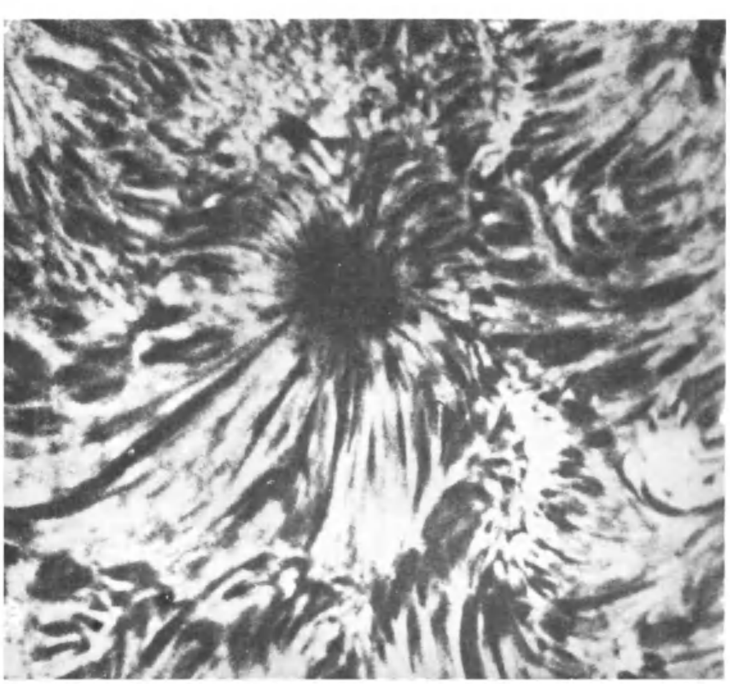
equator. The last spots of a dying cycle appear almost at the equator, but they never appear at the poles.
If you cut off the blinding photosphere with the diaphragm in the focus of the telescope eyepiece, so that only a "black" sun-spot is visible, take care or otherwise you will be blinded. The concept of blackness is relative, and spots that are dark in contrast to the photosphere are themselves blindingly bright.
Spots usually appear in constantly changing groups. New spots emerge, the configuration of old spots changes, but most frequently, the groups break apart and gradually disappear. They may exist for only a few days or for several months.

A spot usually consists of a nucleus, or "umbra", surrounded by a brighter "penumbra" of apparently filamentary structure with traces of vorticity around the centre of the spot. The centre usually lies "slightly" – several hundred kilometres – below the surrounding photosphere. The spots are normally larger than Europe in area and may even be larger than the Earth's total area.

The sun-spots appear dark because the temperature of the solar gases in these places is lower. Judging from the spectrum, it is about 4,500 °C, approximately that of orange stars, whereas the spectrum of the photosphere is the same as that of hotter yellow stars. The lower temperature makes it possible for a great quantity of molecules of chemical compounds to be formed that are only found in small quantities in the photosphere. If a molecule is formed there of some atoms, that have stuck to one another, it will usually break up immediately in violent collisions with other particles that occur at high temperatures.

Appearing at one edge of the Sun and moving across the visible disk, the sun-spots offer a superb possibility for demonstrating the axial rotation of the Sun. The rotational periods determined from different spots and supplemented by spectral determinations (discussed earlier in the book) and other methods show that the Sun does not rotate as a solid body. To a certain degree the rotation of the photosphere can be compared with the movement of tea, which is stirred vigorously in a cup with a spoon. The tea at the centre moves more rapidly than that near the sides. On the Sun the points at the equator make a turn in 25 days; at latitude 60° the period of rotation is 30 days. In astronomic terms the Sun rotates very slowly; at the equator the linear velocity is only 2 kilometres per second.

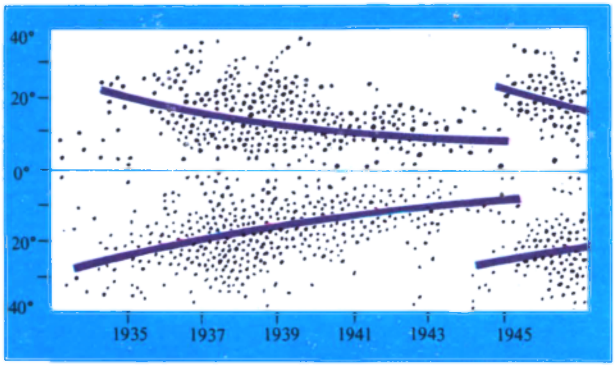
A picture of a part of the solar disk taken in H_α-line showing whirls (1 cm corresponds to about 9,000 km). The black spot in the centre is the size of Eurasia.



The structure of the Sun.

Jupiter and Saturn also rotate in zones like the Sun, more rapidly at the equator, but their period is only 10 hours. For them, like the Sun, the reason for such zonal rotation is unknown, but it is clear that a solid surface could not rotate in this way.

In the Soviet Union and other countries the spots and other phenomena on the Sun are observed on a day-to-day basis. The service is called the solar survey and is conducted in accordance with a common plan by a number of observatories. This effort is necessary to study the nature of the Sun, changes occurring there and their effects upon the Earth. Any amateur who has even a small telescope can observe sun-spots and their changes.

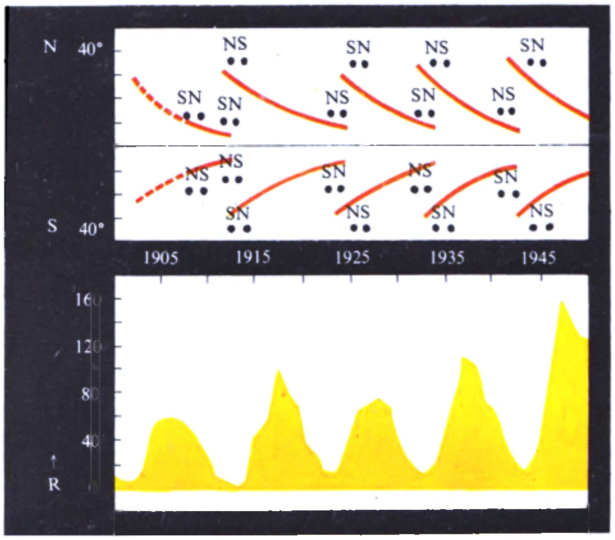


The butterfly diagram of sun-spots.

Observation of the Invisible and the Sun's Anatomy

Astronomers are people who not only know a great deal about things where the untrained eye sees only a scintillating point, but who also seek to observe the invisible. One of the numerous examples of this is their observation of the magnetic character of sun-spots and the distribution of the chemical elements over height above the photosphere.

The polarity of sun-spots: the migration of the spot zone and change of polarity over heliographic latitude (above); and the yearly average number of solar spots *R*.



The spectral lines of a light source change in a magnetic field. They are split, each into several lines, and the light of each of them is polarized in a particular way. Without entering into a long discussion of the polarization phenomenon, we will only say that polarized light can be distinguished from ordinary light by special methods developed by physicists. The spacing between spectral lines into which the initial line is split in a magnetic field increases with magnetic field strength. This phenomenon has been studied thoroughly in laboratories. The dark lines in the spectrum of sun-spots show a similar splitting, so indicating that in the region of a sun-spot there is a magnetic field whose strength sometimes attains 8,000 gauss. This is an extremely strong field, although in laboratories electromagnets can create a still stronger field.

Everyone knows that any magnet has two poles – north and south. In magnets in the form of a bar or horseshoe they are usually painted different colours – red and blue. And here we have a curious phenomenon: sun-spots most frequently appear in pairs, the magnetism of one spot being north and of the other south.

The two spots are like the two ends of a horseshoe magnet lying underneath the Sun's surface with the ends showing through it. Moreover, in all the spot pairs of one solar hemisphere the leading spot (in the direction of solar rotation) always has the same magnetism (say, southern), while in the other solar hemisphere the magnetism of each leading spot is opposite (northern). This lasts for 11 years and when

a new sun-spot cycle begins the magnetism of the spots of the northern and southern hemispheres of the Sun reverses.

In 1958 Babcock of the USA concluded that the general, although weak, magnetic field of the Sun changes its direction. For example, if in the course of an 11-year cycle the north magnetic pole was in the Sun's northern hemisphere, in the following cycle it will be in the southern hemisphere. Why this occurs is still unclear, but the lines of force of the Sun's general field enter the other polar region. The lines of force close, passing within the Sun, but not very deeply beneath the photosphere.

We see the Sun and all the details of its surface in a group of rays of different wavelengths. The solar surface emits a continuous spectrum. But the cooler and thinner layers lying above it, as a result of the scattering processes described earlier in the book, cause the appearance in the solar spectrum of dark lines, which have been named in honour of the German scientist Fraunhofer, who was the first to study them. We have already said that dark spectral lines are not absolutely black, there is a certain amount of light in them. This light is much less than that emitted in the continuous spectrum at this same wavelength and is a tiny fraction of the total light in the continuous spectrum.

If our eyes lost sensitivity to all wavelengths except one corresponding to a spectral line of a particular chemical element, such as hydrogen, the Sun would then appear to us absolutely different from what it is now. In those places where there is much more cool hydrogen above the solar surface, the absorption of light at our wavelength would be particularly strong. We would see a dark spot there. However, where the hydrogen above the Sun's surface is hotter, the light emission at our wavelength would be stronger than at adjacent places and we would see a bright spot there. Thus, it would become possible for us to see at once the distribution of hot and cool hydrogen masses above the Sun's surface.

Exactly such an opportunity to see the Sun "in the light of the wavelength of the hydrogen line" is afforded to us by an instrument known as the *spectrohelioscope*, invented by Hale of the USA in 1930. The spectrohelioscope is a sort of spectroscope in which the entire spectrum is blocked off by a screen with the slit, through which only passes the light of the one desired "dark" spectral line. Behind this slit is an eyepiece for the observer. His eye only perceives the light with the wavelength of the line separated out by the slit in the screen. The spectroscope slit receives the solar image as perceived through the telescope and brought by a special device into a rapid oscillatory motion across the slit. The patterns of rapidly alternating narrow bands, cut off from the solar image by the spectroscope slit, following one after another, as a result of retention of the visual sensation, give the impression of a complete picture of the solar disk. By separating the different lines in the spectrum by means of the slit in the screen it is possible to study the distribution of different gases over the solar surface: hydrogen, helium, sodium, calcium, and others.

In another form of the instrument, called the *spectroheliograph* and invented before the spectrohelioscope by Deslandres of France and by Hale, such images of the Sun can be photographed. Instead of the eye behind the slit in this instrument we have a moving photographic plate. Photographs produced are called *spectroheliograms*.

Theory shows that in dark spectral lines, actually having a finite width (and by no means infinitesimally narrow), the centre of line is due to absorption by gases situated at a greater height above the Sun's surface than the gases responsible for light absorption at the margins of the line. Thus, by separating different parts of broad dark lines of the solar spectrum using a very narrow slit it is possible to obtain, as it were, sections of the gaseous layers at different heights above the photosphere. This is a sheer anatomy of the external parts of the Sun.

On the spectroheliograms it is possible to see clearly the structure of hydrogen masses in the region of spots, which, as mentioned above, is not visible in an ordinary telescope. In addition, the spots are normally surrounded by bright clouds of hot hydrogen and calcium (flocculi). Flocculi are the upper parts of the regions occupied by faculae. The fact that the emergence of cooler regions is accompanied by clouds of hot gases has the result that when cool spots occupy a maximum area the total solar radiation apparently does not decrease.

Examination of spectroheliograms to measure the velocities of gases at different places on the Sun shows that gases at spots circulate in a complex manner.

In the lower part of a spot the gas flows radially from the centre and in the higher layers it flows into the spot from the outside. The velocities may be as high as 10 kilometres per second. However, spots are quiet formations where convection is suppressed by a strong magnetic field. But around the spots, in the region of flocculi, the magnetic field is weak and the convection of ionized gas, called plasma, is more intensive.

Solar gases are involved in continuous and powerful circulation, whose laws are for us far "darker" than the sun-spots themselves.

The Envelopes of the Sun

Although the photosphere itself consists of thin gases, it is surrounded by an atmosphere, which is still thinner. It would perhaps be better to say that the Sun is surrounded by several envelopes, as if inlaid one into another, so that the solar atmosphere consists of several layers. The solar atmosphere, more tenuous than the photosphere, is almost completely transparent. We see the photosphere through it, as if through clear glass, but we do not see the atmosphere itself, like glass. The solar atmosphere is heated to several thousand degrees and therefore emits light.

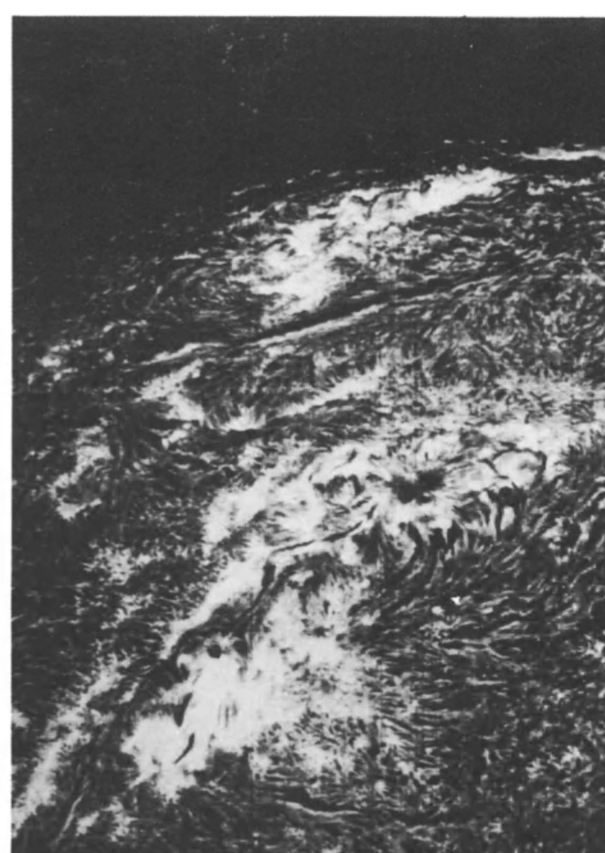
The relatively thin and tenuous layer of the atmosphere emits not a continuous spectrum, but bright lines, although so weak that against the bright background of the sky near the solar limb the atmosphere is invisible. The sky, illuminated by the Sun near its limb, is brighter than the solar atmosphere. Where the photosphere can be seen through the atmosphere the light of the latter is absorbed, in accordance with principles with which we already are familiar, at wavelengths that the atmosphere emits.

The absorption of light by the solar atmosphere at particular wavelengths also causes the appearance of dark Fraunhofer lines in the continuous spectrum of the photosphere.

But if the atmosphere of the Sun, projected on its disk, is invisible at the limb against the bright sky, can we determine the structure of the atmosphere?

Yes, we can. This can be done during total solar eclipses. When the Moon blots out the blindingly bright solar disk, the sky near the Sun, no longer illuminated by its direct rays, darkens. Then the solar atmosphere becomes visible against the darkened background of the sky from behind the lunar disk. The solar atmosphere appears as a bright ring framing the dark circle of the Moon. Let only a tiny piece of the bright photosphere appear from behind the Moon, the sky immediately becomes bright and again the solar atmosphere becomes invisible. At a total eclipse the solar atmosphere can be observed only a few minutes, no more. In addition, total solar eclipses visible in areas where it is convenient to send teams are rare, so that on the whole we can only see the solar atmosphere

An H_α-line picture of the solar disk showing the structure of the chromosphere.



in such a way for about 1-2 hours–and this since the time when science became interested in the phenomenon.

The term solar atmosphere is usually given to the layer of gases that are more tenuous than the gases of the photosphere, decreasing in density with distance. This layer of varying thickness at the time of total eclipses appears from behind the dark Moon either as a reddish ring or as a crescent, depending on the circumstances and phase of the eclipse. Owing to its reddish colour this envelope of the Sun has been called the *chromosphere*, the reddish colour being due to emission by hydrogen, the most important component of the chromosphere. The spectrum of the chromosphere consists of bright lines on a dark background. Its lower parts contain all those gases that are responsible for their absorption in the denser outer layers, thereby producing dark lines in its spectrum. The spectrum of the lower parts of the chromosphere, consisting of a great many bright lines, is visible for a short time, usually 2 or 3 seconds, and so it has been called a flash spectrum, its lines being visible momentarily. The different chemical elements making up the chromosphere are observed to different heights. The highest one (up to 14,000 kilometres) is ionized calcium, although it is heavier than hydrogen.

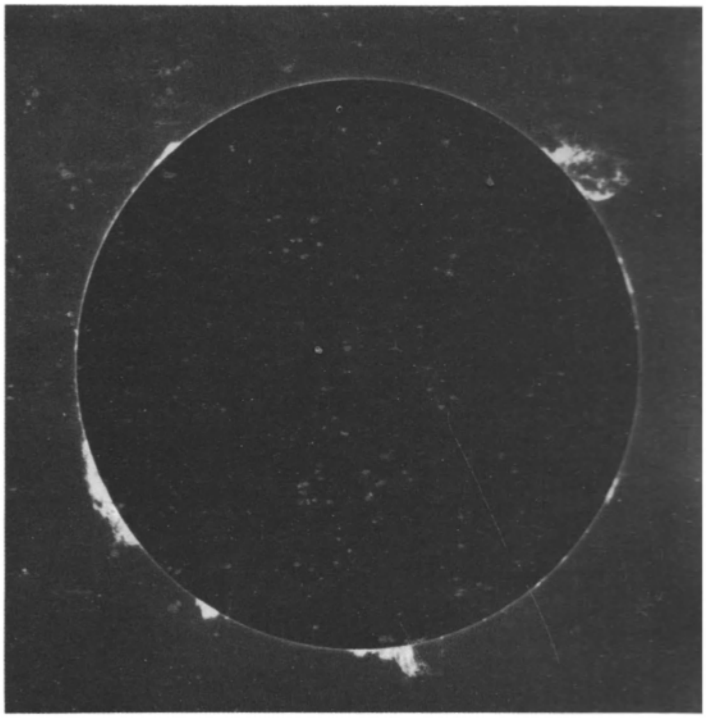
The apparent boundary of the chromosphere, different for different gases, also varies constantly because it is apparently not a fixed, unperturbed superposition of layers of gases, but is rather a result of surges of gases in the form of countless streams or fountains rising from the reversing layer or even from the photosphere. Just imagine a countless number of small fountains whose jets merge into a continuous wall of water–this is the approximate model of the chromosphere. In other words, the chromosphere is not a static, but a dynamic formation.

In any case, the presence of strong vertical movements of gases within and from the chromosphere and even explosion-like surges of gas are observed constantly.

The Highest Fountains in the Universe

During total solar eclipses it is possible to see, even with the naked eye, gigantic fountains of hot gas surging from the atmosphere, these are called *prominences*. A first such prominence recorded in history was observed in ancient Russia in the year 1185, but it was not until many centuries later that the physical nature of prominences was clarified. Gases surge with velocities attaining hundreds of kilometres per second, but however enormous these velocities are, they are usually inadequate for detaching the prominence from the Sun. On the Sun's surface the critical velocity, at which a body can overcome solar attraction and escape to infinity, is 617 kilometres per second. The gases of prominences, rapidly rising upwards, expand to form extensive clouds and, dispersing, precipitate downwards. Prominences can reach colossal height. For example, in 1928 astronomers observed a prominence that attained a height of 900,000 kilometres, i. e. 1.3 solar radii. This is 2.5 times larger than the distance from the Earth to the Moon. A prominence was observed in 1946 that reached twice this height. With the enormous velocity of the eruption the changes in prominences occur very rapidly, literally before our eyes.

In addition to such prominences, which are called



Solar prominences of 29 August 1958.

eruptive prominences, consisting almost entirely of the gases of which the chromosphere is made up, it is possible to observe at the limb a type called quiescent prominences. They have the form of enormous clouds floating over the chromosphere and connected to it by individual columns or branches. They emit the lines of hydrogen, ionized calcium, and helium. They are sometimes up to 600,000 kilometres long – 50 times the Earth's diameter – and yet such a prominence is only a small temporary appendage of the solar atmosphere.

The average prominence floats at such a height over the surface of the chromosphere that the Earth could freely pass under it. With a mean thickness equal to the diameter of the Earth, a length of 20,000 kilometres and a height of 50,000 kilometres, a prominence has a volume 100 times that of the Earth, but since it consists of thin gases, its mass is only 10^{18} grammes, or the mass of a cube of water 15 kilometres on the side, which is still larger than the mass of a small asteroid.

The number of solar prominences changes from day to day, but for the most part it goes up and down in accordance with the area occupied by sun-spots. In addition, eruptive prominences arise near sun-spots, whereas quiescent prominences are encountered anywhere on the Sun's surface.

Prominences can be seen "in profile" at the solar limb during total eclipses, but they can also be seen "from above", in projection on the solar disk. Having a temperature of about 10,000°, like the chromosphere, they absorb the light emitted by the photosphere at wavelengths corresponding to the absorption power of the atoms of which they consist. Therefore, on the spectroheliograms they appear as long dark filaments. Using a special method, to be described below, it is possible to observe prominences at the solar limb constantly. Thus, we can observe prominences daily over the entire solar disk.

Prominences seem to be held at a great height by electromagnetic forces, but they vary in size from place to place and even in the same place due to some physical processes, sometimes even in jumps. The velocities of motion of prominences, as has been discovered recently, sometimes change in jumps as well.

Stars, just like the Sun, should have a photosphere and atmosphere consisting of a reversing layer and chromosphere. It is their difference in temperature, composition and structure, if any, from the conditions on the Sun that governs differences between the solar and stellar spectra. Stars should also have prominences, although it is impossible to observe them directly.

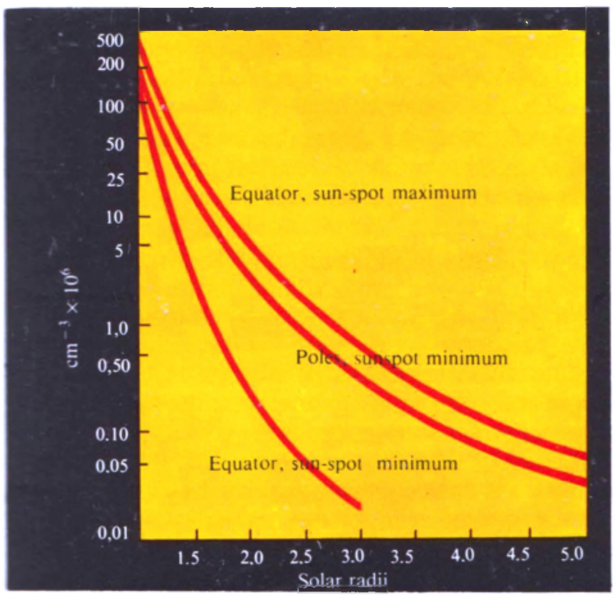
Lyot in France and specialists at Harvard Observatory in the USA and the Crimean Observatory in the USSR used special light filters that only transmit the red hydrogen spectral line emitted by prominences to photograph them on movie film. The film can be run rapidly, showing how these gigantic gaseous arches rise from the Sun's surface, and then are dispersed. The films also show that at a certain height above the Sun's surface luminescence of the prominence appears suddenly and how it then propagates towards the solar surface, rather than away from it. Especially attractive are previously unknown prominences that rise obliquely over the Sun's surface, like the water from a hose, and then fall back to the Sun's surface. They travel back along the same path, pretty much like a stretched worm that suddenly contracts. These remarkable photographs shed new light on the nature of prominences, suggesting the involvement of electromagnetic forces in the process of their changes.

The Solar Corona and Its Mysteries

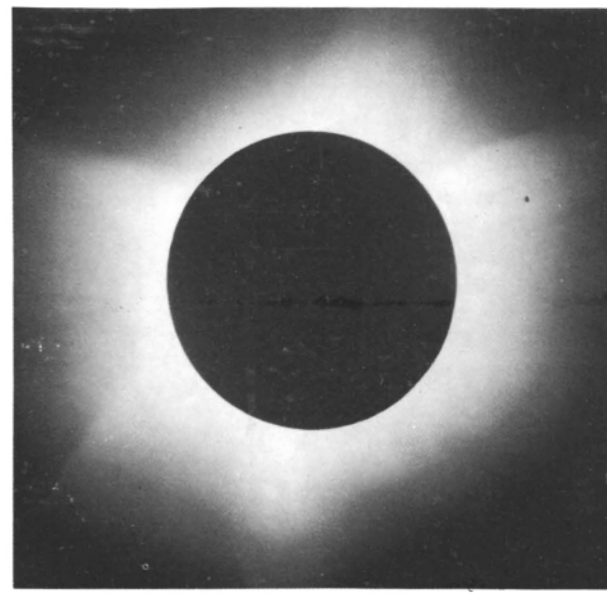
The Sun is frequently called the lord of the solar system. This analogy is not entirely apt, although the lord of the planets actually is crowned by a remarkable crown, or corona, and this corona is the one of pearls. Strictly speaking, it is its colour that comes from pearls, but as to what the corona is made of, it will be the subject of our further discussion in considerable detail.

Only during total solar eclipses can we see the solar corona as a remarkable silver-pearly radiant glow around the Sun. The inner part of the corona, which is the brightest, gives a continuous spectrum on which there are superimposed bright lines, no one of which has ever been observed in laboratory on Earth. The outer part of the corona, which is less bright, is characterized by streamers, which may be as long as the diameter of the Sun or even longer. The total light of the corona is approximately half that of the full Moon.

The corona's shape differs from eclipse to eclipse. The Pulkovo astronomer Gansky, who passed away



The electron density in the solar corona.



The solar corona (the eclipse expedition, Takengon, northern Sumatra, 9 May 1929).

prematurely in 1908, discovered that this shape is dependent on the phase of solar activity.

When there are many spots and prominences on the Sun the corona has a rugged appearance. Its curved streamers protrude in all directions, like the hair on a man's head when he has just awoke. When there are few spots on the Sun, however, the corona extends along the solar equator like wings or fans.

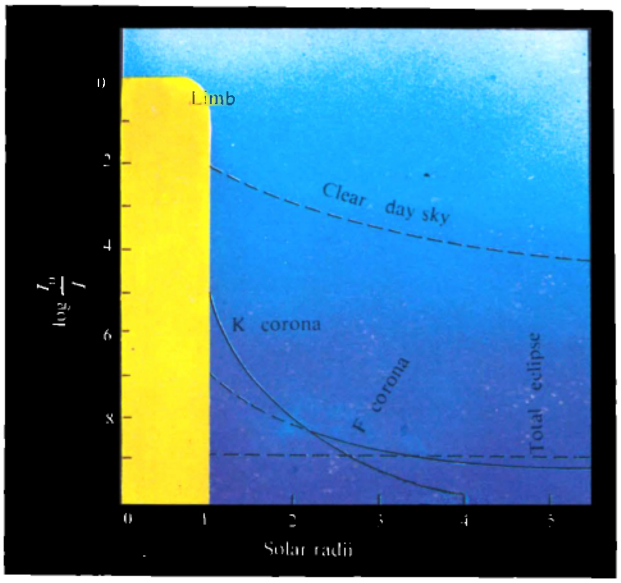
In 1942 the Soviet astronomer N. M. Subbotina advanced the interesting hypothesis that the remarkable image of the winged Sun of the Egyptians, their most sacred and beloved symbol, equally with the scarab, is none other than the image of the Sun and its corona.

In any case, the corona, which is clearly visible with the naked eye at the time of an eclipse, must have produced a tremendous impression on the observing Egyptian priests, who deified the Sun and devised the image of the winged Sun.

On some of the pictures of eclipses in epochs intermediate between the sun-spot maximum and minimum, the solar corona looks like the wings of a gigantic butterfly that flew up into the sky and landed on its violet-blue velvet background.

Several thousand years ago the constructors of Egyptian pyramids gazed at the remarkable and mysterious phenomenon of the corona, on the winged Sun, but we, too, must admit that for us it still presents a good many mysteries.

We still do not know the origins of the corona and the causes of its radiation, although the curvature of the coronal streamers from the poles towards the solar equator suggests the idea of lines of force of the magnetic field of a magnetized sphere. The corona is not a quiescent, static formation, it is supplied continuously by matter coming from the Sun or by matter flowing in from the outside. The immense extent of the corona and its spectrum (not purely gaseous) do not enable us to call the corona the outermost part of the solar atmosphere. If it is formed by meteoric dust flowing towards the Sun, as certain scientists think, it cannot be called an atmosphere, but if it consists of matter scattered around by the Sun itself, it constitutes a singular and grandiose atmosphere. In any case this should be true of the inner gaseous corona, since it is adjacent to the gases of the solar atmosphere and forms a relatively thin layer above the latter.



The sky brightness near the solar disk.

The outer corona gives a spectrum that is a replica of the solar spectrum – continuous and with the same dark lines. It is postulated that the outer corona consists of electrons, but farther from the Sun – of hard particles of meteoric dust, which scatter sunlight. Fesenkov has pointed out that the outer corona, in that part of it consisting of meteoric dust approaching the Sun, need not be regarded as reaching down almost to the surface of the Sun itself. It can cut off at a distance of about 0.1 AU from the Sun, since the closer meteoric particles will already be evaporated. However, this does not prevent us from viewing the corona as a halo surrounding the Sun and slowly increasing in brightness with apparent approach to the solar surface in projection on the celestial sphere.

Still another mystery is the spectrum of the inner corona, which consists of bright lines. But we will discuss this below.

The radio emission of the outer corona is too weak to be measured. But the electron concentration there is sufficient for radio waves to refract in it. Vitkevich suggested observing annual occurrences of the Crab Nebula, a very powerful radio emitter to be discussed in the section "Supernova Explosions", when it is blotted by this corona as the Sun moves along its ecliptic. The refraction temporarily changes the "apparent" position and configuration of the Crab Nebula, as "seen" through the radio telescope. It was thus that the outer corona was uncovered, which by 1967 was observed over the distance of up to 100 solar radii, which amounts to half the Sun-to-Earth distance.

How Three Astronomers Deceived Nature

This "deception", as well as many other "deceptions", was based on a profound knowledge of this same nature. In our studies, the interferencies created by some properties of nature can be removed by exploiting its other properties.

On 26 October 1868 a just received letter from the British astrophysicist Sir Norman Lockyer was read at a session of the Paris Academy of Sciences. It was dated 20 October and contained a description of a method that Lockyer had devised for the observation of prominences at any time, not solely during total eclipses.

Lockyer used this method successfully for daily observations of what earlier could be observed for only a few minutes at an eclipse, once in several years, and which required distant travels to the zone of the totality.

Those present had not yet recovered from their delight and astonishment at this valuable invention of Lockyer's when the secretary of the Academy of Sciences took a second envelope with the postmark Guntur, India, and read the letter. It had been written by the French scientist Jules Janssen on 19 August, but it had just reached Paris. Janssen reported his discovery of a method for observing prominences daily, when there was no eclipse. The Janssen method was absolutely identical to the technique that had been independently and simultaneously discovered by Lockyer. The two scientists were separated by a distance of a quarter of the Earth's circumference.

Janssen had set out on a distant sea voyage to India to observe the total solar eclipse of 18 August 1868. Pointing the spectroscope at prominences rising above the solar limb, which became visible as soon as the Moon had covered the Sun, Janssen saw that the spectrum of the prominences consists of bright lines.

An unexpected thought flashed in his brain and he immediately cried out to his associates: "I can see

these prominences even when there is no eclipse!" Indeed, on the following morning he showed them to all who wished to see them, despite the fact that the eclipse was long over.

The idea of Janssen and Lockyer involved a magnification of the contrast between the prominences and the sky near the solar limb caused by the differences in their spectra. At most times the viewing of prominences is hindered by the bright background of the sky, illuminated by the Sun.

If the broad slit of the spectroscope is set tangent to the solar image at the focus of the telescope in such a way that the image of the prominence enters it, then the light that comes from the prominence will be broken down into its several bright coloured images according to the emitted wavelengths. However, the skylight, which enters the spectroscope slit as well, is distributed through the entire continuous spectrum because skylight is the scattered light of the Sun and its spectrum is the solar spectrum.

As a result, the brightness of the coloured images of the prominence on the background of the continuous spectrum will be enhanced, its contrast will be greater than in an ordinary observation and the prominences will become visible. Not only will they become visible, but it will be possible to identify the gases of which they consist, since images of prominences in the spectrum are obtained only at wavelengths emitted by the constituent gases.

In memory of this astonishing coincidence the Paris Academy of Sciences struck a gold medallion with the portraits of Janssen and Lockyer on one side. On the reverse side was the image of the Sun god Apollo in a chariot harnessed to four horses, and the inscription *Analysis of Solar Prominences on 18 August 1868.*

But where is the third scientist, you may ask, since in the heading of this section reference was made to three scientists. The third scientist appeared on the scene 60 years later, and this scene was the Midday Peak (Pic du Midi) in the French Alps. Its elevation is 2,800 metres above sea level. But before the scientist attained his success, many others had worked on this problem and all had suffered bitter disappointment.

Everyone wanted to see prominences directly and at the same time all around the Sun, and not in a spectroscope slit. Still others dreamed of seeing or photographing the solar corona when there was no eclipse.

It seemed that after many attempts and failures by scientists of all countries a previously unknown German amateur astronomer, named Blunck, finally solved this problem in 1930. He expected to photograph the corona through a glass that only transmitted infrared light, hoping that in this light the contrast between the light of the corona and the skylight would be greater than in ordinary light. After stubborn work for some years he was able to devise a special kind of plate and publish a picture of the Sun's corona taken at a time when there was no eclipse. Towards the end of his labours the inventor fell seriously ill from inhalation of the vapours of toxic substances he used in his experiments. Alas, his sacrifice was in vain, since it was soon demonstrated that Blunck had not photographed the corona but the halo created around the Sun by the dust particles of our atmosphere and easily seen by the naked eye.

Following the failures of Blunck, success seemed impossible, but in 1930 Lyot described experiments he had performed with a coronograph he had designed. With the use of this instrument installed on mountain peaks, where the sky is darker and clearer than below, it was possible to see the prominences around the Sun on any day and observe the bright lines of the solar corona in the spectroscope, although the corona itself could not be seen.

Lyot decided that the contrast between the background of the sky and prominences, and also the corona, would be improved if in the telescope light would be scattered less, since it was impossible to any adequate degree to decrease its scattering in the Earth's atmosphere. The influence of the latter could only be attenuated by ascending a mountain, leaving behind the most strongly scattered light.

The folowing fascinating experiment will show the extent to which the contrast between a source and the observer decreases as a result of light scattering. Take a box, and insert a slide showing the solar corona against a dark background in place of one of its walls. Scrape out the place occupied by the black disk of the Moon on the slide and insert an electric bulb into the box. Suspend a piece of gauze in front of the box. The fabric will, like the atmosphere, scatter the light of the lamp passing through it to the observer. Now gradually cover the transparent circle illuminated from within by the lamp, which will represent the Sun. As long as even a single ray of the lamp (Sun) falls on the

gauze (atmosphere) the corona is invisible. As soon as the circle of the Sun is covered completely by a suitable opaque cardboard circle, the faint glow of the corona will immediately flare up around the "eclipsed Sun". Suspend a coarser gauze, which causes less scattering, and the corona will appear still more clearly.

In order to decrease the light scattering in the telescope Lyot made an objective of the most transparent glass, protected it against the tiniest of scratches and dust particles and eliminated each dust particle from the air inside the telescope. Each of these details, combined with the others, appreciably reduced light scattering on the path from the Sun to the observer's eye: many a little makes a mickle. Placing a black circle in the telescope's focus, a circle that just barely covered the solar image, Lyot could perceive the rose-coloured prominences directly in the eye-piece of his coronograph. The three scientists thus outwitted nature, which was preventing us from studying prominences and the corona. Now Lyot coronographs are widely used in solar studies throughout the world.

Solar Chemistry

The spectrum of the lower chromosphere, observed for one or two seconds during total eclipses (which is why it is called the flash spectrum), and the dark Fraunhofer lines in the ordinary solar spectrum make it possible to determine the chemical composition of the solar atmosphere. This must be kept firmly in mind. It is, however, impossible to determine the chemical composition of the Sun's interior from the spectrum: we only see the spectrum of the atmosphere.

We only find those elements on the Sun that are known to us on Earth, but not all of them. Of the 92* elements of the Periodic System a total of 67, or two-thirds, have been detected in the solar atmosphere.

In any former book on astronomy you will read that the Sun, the "golden Sun", contains no gold. But from this one you will learn that there is gold on the Sun, although an insignificant quantity. It was discovered in 1942 from an extremely weak line whose origin had earlier been unclear.

The chemical elements absent in the solar spectrum may of course be actually absent in the solar atmosphere, but there can also be other reasons for the absence of noticeable lines in the spectrum. For example, this can be caused by the small content of this element, an actual absence of its lines in the part of the spectrum accessible to observations, or an inadequate study of the spectrum of this element in the laboratory.

The solar spectrum lacks the lines of the majority of the heavy radioactive elements, rare earths, inert gases (except helium and neon) and halogens, but it does have the lines for technetium, a product of nuclear reactions.

In addition to the atoms of many elements, in the solar atmosphere, predominantly in the region of spots (having a lower temperature) very simple molecules have been found: carbon, cyanogen, hydrogen compounds and many others. Moreover, within the spots there are titanium oxide, magnesium, aluminium and calcium hydrides, aluminium and zirconium oxides, and other compounds.

Study of the intensity of the solar spectral lines makes it possible not only to determine different elements in the solar atmosphere, but also to work out their quantities. Thus, it has been established that the solar atmosphere contains:

	Per cent, by volume	Per cent, number of atoms
Hydrogen	81.760	90.7
Helium	18.170	9.1
Oxygen	0.03	0.09
Magnesium	0.02	—
Nitrogen	0.01	0.01
Silicon	0.006	—
Carbon	0.003	0.05
Iron	0.0008	0.007
Calcium	0.0003	less than 0.01
Neon	—	0.01

It follows that the Sun is more than 70 per cent hydrogen and 28 per cent helium by mass. It has been estimated, however, that the solar *interior* is poorer in hydrogen.

* 107, including the artificially produced elements heavier than uranium.

A History of Two Strangers

In 1868 astronomers turned their attention to the fact that the spectrum of prominences contains a bright yellow line which had never been observed previously.

It obviously belonged to some matter that did not exist on Earth and was only present on the Sun. It was proposed that this element be called *helium*, from the Greek word *helios*, which means "sun". Astronomers postulated that this must be a very light gas because it rose high in the Sun's atmosphere. During twenty-five years it was assumed that the Sun contained a unique "solar element", helium.

In 1893, during a new precise determination of the mass of nitrogen, the British physicist Rayleigh discovered a discrepancy between the mass of nitrogen, produced from ammonia and from the air. In a litre of gas the disagreement was equal to the mass of a flea. But Rayleigh was not content with this situation and began to search for the cause. In order to solve the problem more quickly he invited the collaboration of the famed chemist Sir Ramsay who suspected that the nitrogen produced from the air was not pure and contained some gas heavier than nitrogen. As a result, the mass of the nitrogen produced from the air was larger, just as dirty salt with an admixture of sand is heavier than the same volume of pure salt.

Cudgeling his brains over his problem, Ramsay recalled the description of one of Cavendish's experiments, something he had read as a student. Cavendish made his experiment in 1785, but it was overlooked. Cavendish had used electrical discharges to combine nitrogen and oxygen, resulting in nitric oxides. Try as he would, in the container with mercury that displaced the nitrogen that was there earlier but had become an oxide, there was still a tiny bubble of gas, which by no means wanted to combine with the oxygen. Cavendish took his nitrogen from the air, and so Ramsay suspected that Cavendish had come across the gas that was causing so much concern to his friend Rayleigh. He set out to find the identity of this air bubble. He did so by repeating the Cavendish experiment on a larger scale and he obtained a stubborn gas that did not desire to combine with oxygen. The gas was no longer in a bubble, but in a volume that made it possible to make precise determinations of its mass. He experimented in another way, by passing nitrogen produced from the air through red-hot magnesium until they combined completely. The residue was the same inert gas that did not want to combine. It was 1.5 times heavier than nitrogen.

The new gas did not want to combine with any other element either. Because of its chemical "idleness" it was called *argon*, a Greek word for "idle". Argon was a new chemical element, and it would have been discovered 100 or more years earlier if Cavendish, who "held it in his hands", had had precise scales to weigh his bubble of gas.

Rayleigh and Ramsay rested content with their discovery but their rest was not to be for long. A chemist wrote in 1895 that the famed explorer Nordenskjöld had long ago brought a new mineral from Norway – cleveite. From this black mineral it was possible to separate out a gas that would not combine with oxygen. "The geologist who has described this mineral assumes it to be nitrogen, but in actuality it may not be nitrogen, but argon," wrote the chemist.

Ramsay then acquired some cleveite, separated out the gas and studied its spectrum. The spectrum was not at all that of argon. This was something new, with a bright yellow line. He recalled that astronomers had discovered such a line a quarter of a century earlier in the spectrum of prominences and had observed it daily since, attributing it to the nonterrestrial gas helium. Thus, a solar element – helium – was discovered on Earth.

The astronomers were right. Helium is a light gas, the second lightest after hydrogen. Small quantities of helium have also been discovered in the air. Helium features a great many interesting properties, in particular, it goes liquid at a very low temperature (– 269°C). Interestingly, three other unknown gases were discovered as helium was investigated.

In 1914, British artillery unsuccessfully cannonaded a German zeppelin that was heading for Paris. Although the craft was hit by shell splinters it did not burst into flames, like all hydrogen-filled dirigibles. British chemists figured out that the Germans filled their zeppelin with helium, but it was a mystery where they had obtained it. At that time only a small quantity of helium, produced with great

difficulty, was in the hands of scientists. The British government undertook urgent measures to find natural helium in their possessions and in 1918 it was found in petroleum gases in Canada, whence it began to be extracted for military purposes. It was not until 1930 that the British had accumulated sufficient helium for the filling of the dirigible R-100. It was found that Germany had produced helium from monazitic sand, which it had imported aboard steamships from India and Brazil as ballast.

There is an extraordinary amount of helium on the Sun, but of course it is inaccessible to us. Forty years after helium was discovered on Earth it still was an extreme rarity, but now electrical discharge in this solar gas sparkles on Earth in a rosy-yellow colour in the shop windows of stores and ads. The solar element has become a matter of course.

The unmasking of a mysterious stranger that appeared before us in the form of a number of coloured spectral lines in the solar corona has had a different history. The stranger, discovered in 1869, a year after the discovery of helium, was called *coronium*. Coronium was more stubborn than helium: it did not want to appear on Earth. It still has not shown up. Moreover, helium at least has been discovered in the spectra of other celestial bodies. stars and nebulae, but coronium has only been discovered in the solar corona.

It was only in 1933 that coronium appeared for several months in the spectrum of the so-called novalike star RS Oph in the Serpent Holder. In that same year a faint star, which had already flared up relatively recently, in 1893, flared up again and coronium was observed briefly in the gases it ejected at its maximum brightness. This case was of interest to astronomers, but it shed no light on the secret of coronium, although certain indications even then led the author of this book to the thought that coronium is somehow associated with iron.

The fact is that each chemical element in the Periodic Table has its own pigeon-hole where it fits. As soon as a new element is discovered, it is immediately assigned to an appropriate hole in a particular row of the table. Almost a quarter of a century back virtually all the holes for the elements in the table were occupied. There was no place in the table for new elements. This means also that there are no places in nature for new elements. Hence any new stranger, including coronium, is not a stranger at all, but some old acquaintance in a fancy dress and mask. The name "coronium" given to it is not its real name, but a pseudonym under which it is hiding. Its masquerade is forced and its fancy dress is unfamiliar spectral lines it acquired when it got into the unusual physical conditions prevailing in the solar corona, conditions not present on Earth. Tear away the mask and you will see beneath it the familiar oxygen, nitrogen or some other element, laughing at our futile efforts to unmask it.

The conviction that it was possible to identify coronium in such a way was supported by the success with the spectrum of tenuous masses of gas forming nebulae in interstellar space. The element discovered in them, called *nebulium* and hidden under green spectral lines, was subjected to "trial by ordeal" by the US physicist Bowen. After a long quest he identified it in 1927 as simply oxygen. However, it was not simply oxygen, but oxygen that was doubly ionized, i.e. had lost two electrons. But we would have been able to unmask it earlier if it had not contrived to emit spectral lines that it was "forbidden" to emit.

Strictly speaking, nothing prevents the emission of these lines, but their emission under terrestrial conditions is difficult, it is virtually impossible to detect them in the oxygen spectrum on Earth, and physicists therefore have arbitrarily called these lines "forbidden". The lines appear when an electron jumps, without being forced, from one orbit to the next. But in the orbit from which it should jump back, having spun in it for a certain prescribed time, the electron remains very long – seconds, hours, days and even months – before the electron of itself abandons the orbit and emits the corresponding "forbidden" line.

Under terrestrial conditions the densities of gas are so great, and hence the collisions of atoms are so frequent, that with such an orbit an electron in a collision is driven forcefully into another orbit sooner than it would have gone there of its own free will. In ordinary orbits an electron remains only about 10^{-8} seconds. This does not make it possible for an atom to emit a forbidden line.

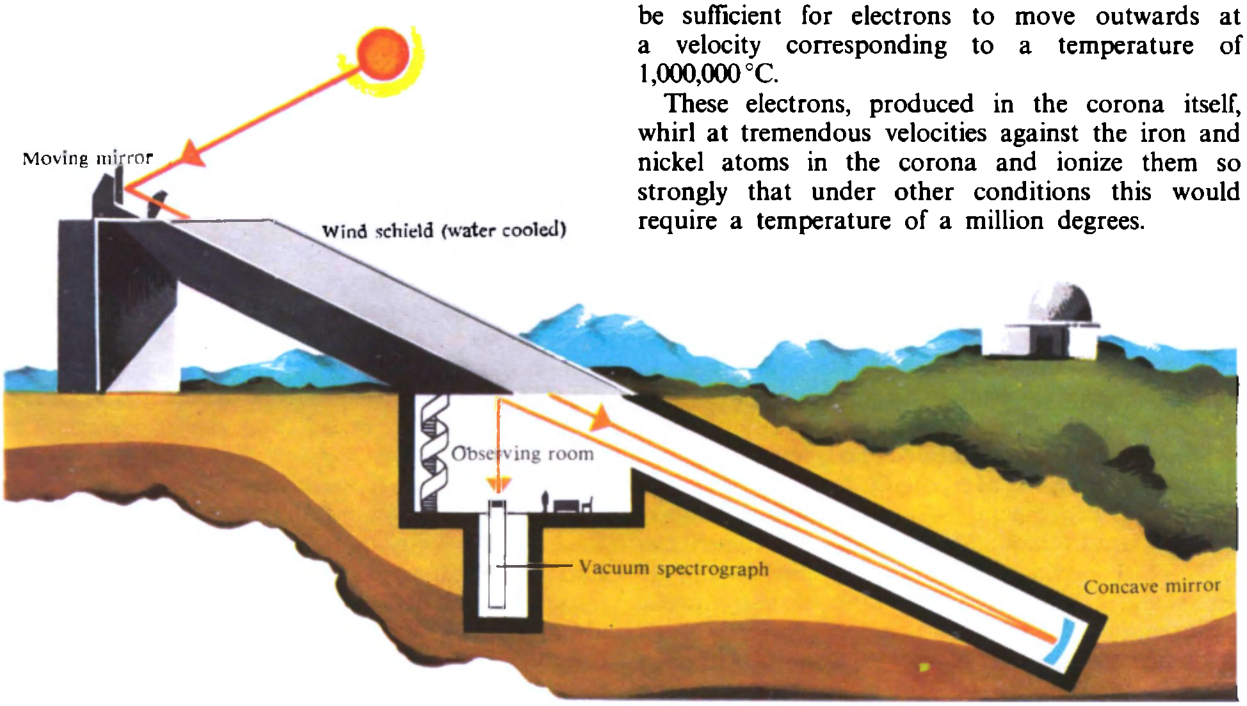
In gaseous nebulae the gas density is so negligibly small that collisions of atoms are extremely rare and

the lines they emit are forbidden under terrestrial conditions. In nebulae, however, they occur without hindrance.

Like with nebulium, the lines of coronium were sought first among the forbidden lines of the known elements. Their wavelengths can only be determined theoretically, knowing the structure of atoms, but this structure is not yet known for all of them. In the gases ejected into space by the star RS Oph there were unusually intense forbidden lines of iron atoms, not too strongly ionized. It was therefore possible to refer the coronium lines in the spectrum of this star to the unusual conditions for the luminescence of iron vapours characteristic of stars. The many attempts to do so were fruitless, but in 1941 the Sweden scientist Edlén reported the long-awaited news – "coronium is iron".

Some coronium lines were forbidden lines of nine-fold ionized iron and others were the same lines of thirteen-fold ionized iron, the fainter lines belonged to multiply ionized nickel and other elements.

The diagram of the Kitt-Peak solar telescope (Arizona, USA).



The gas density in the corona is undoubtedly very small and could permit the emission of forbidden lines. But the iron in the solar corona could be the result of evaporation of iron meteor dust when this matter approached sufficiently close to the Sun.

At first astronomers were a bit suspicious of this identification of coronium. How was it possible that so close to the Sun, which has a temperature of "only" 6,000 °C, could there be iron ionized so strongly? Such ionization under ordinary conditions requires a temperature of above 100,000 °C and therefore no one earlier had sought coronium among ions whose existence requires such high temperatures.

Recently, however, it became known that in the neighbourhood of the Sun there are iron vapours whose atoms are stripped of 9 and even 13 electrons. This can occur not only due to high temperatures, but also due to other processes proceeding in different places in the chromosphere. Their description is beyond the scope of the book, but we will note that the Moscow astronomer Shklovsky gave the following reasoning: under the conditions prevailing in the corona a weak electric field would be sufficient for electrons to move outwards at a velocity corresponding to a temperature of 1,000,000 °C.

These electrons, produced in the corona itself, whirl at tremendous velocities against the iron and nickel atoms in the corona and ionize them so strongly that under other conditions this would require a temperature of a million degrees.

It was shown that the surface of the Sun, by casting its electrons into outer space, becomes positively charged due to the accumulation of positively charged ions. As a result, ions are mutually repulsed and ejected from the Sun, whose charge decreases so enabling electrons to be ejected again. Thus, the Sun gradually loses its mass.

There is no more mystery of coronium, just as there is no more mystery of the spectrum of the Sun's corona as a whole. Now there does not remain in the spectra of celestial bodies even a single "celestial element"; they all appear to belong to elements occurring on Earth. We have lost in science two chemical elements - nebulium and coronium - but instead we have acquired knowledge concerning the structure and behaviour of both the tiniest of atoms and of enormous bodies in the Universe...

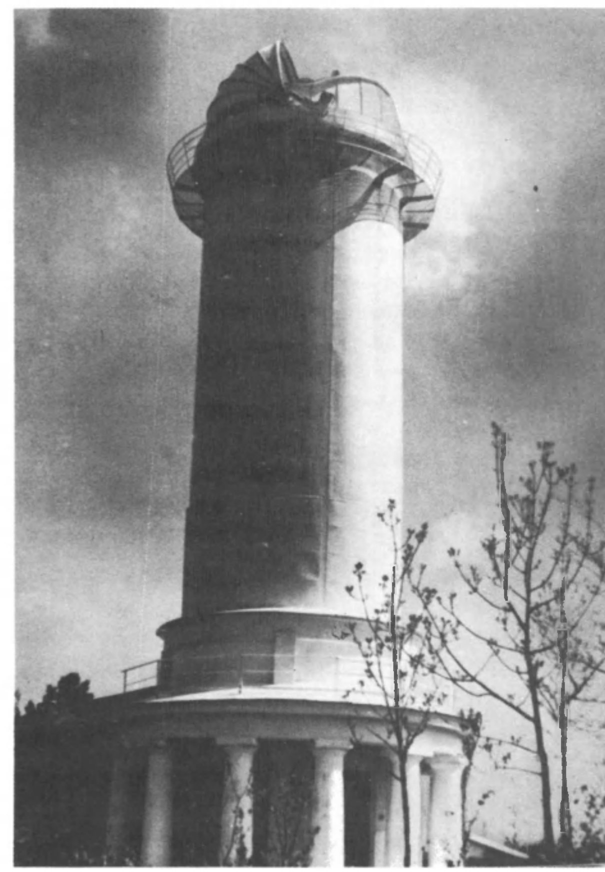
Active Regions, Chromospheric Flares, X-Ray and Radio Emission of the Sun

The so-called active regions of the Sun are characterized by more intensive motion of gases. Not only spots emerge there, but also faculae, flocculi, stronger magnetic fields, and some prominences.

Active regions produce more corpuscular, ultraviolet, X-ray and even cosmic-ray emission of high energy. It was not until fairly recently that these radiations could be studied by instruments installed on satellites, high-altitude rockets and planetary probes. For example, the ultraviolet end of the solar spectrum is strongly absorbed by the terrestrial atmosphere. But it contains valuable information about the physical state and chemical composition of the outer layers of the Sun.

X-ray solar radiation is generally measured using counters with a thin film that absorbs X-ray quanta of different energy depending on the composition and thickness of the film.

Spectroheliograms indicated that now and then there are bright flares of especially dense and hot gases on the Sun, the temperature of the plasma in the flares reaching tens of thousands of degrees. It is these flares and not the spots with which they are commonly associated that account for electromagnetic perturbations on Earth, which have formerly been attributed to the direct effect of the spots. (Although the picture seems to be still more complex.) The perturbations manifest themselves in the vibration of the magnetic pointer of a compass, noises in telegraph and radio communications, etc., to be discussed later in the book.



The solar tower at the Crimean Observatory (USSR).

Radio communication would especially benefit from possibility to predict occurrence of such interferences. Experiments are under way to forecast the disturbances, or even weather, based on attempts to liken disturbances to occurrences of active changes and active radiation on the Sun. The fact is that, generally speaking, in order to be able to affect the Earth an active region must lie near the centre of the apparent solar disk. Knowing the solar period it is always possible to predict when the active region that is now far away from the centre would come to it (or rather to its central meridian).

Improved predictions of chromospheric flares are important to provide the safety of astronauts. The flares produce rays similar to cosmic rays in composition: 90 per cent of protons and 10 per cent of alpha-particles (helium nuclei). In the process the cosmic rays may increase in intensity thousand-fold over several hours. The largest flares occur on average once every 4-5 years during maximum or minimum sun-spot periods.

In 1957 attempts were undertaken to find on the Sun the hydrogen isotope with atomic mass 2 (deuterium). It was expected that it might form in nuclear reactions during the solar flares. In August 1972, during a large flare gamma-radiation was detected that might be produced as deuterium was formed. Next year satellite-borne instruments found deuterium directly in the solar wind produced by several flares. Another hydrogen isotope – tritium – was found too. It is unstable and a half of it decays in 12.6 years. Both isotopes are produced in collisions of fast protons and helium nuclei with nuclei of heavier elements. In flares deuterium increases hundreds of times, reaching 0.1 per cent of hydrogen atoms. The intensely active region on the Sun, which seems to have occurred at the time, gave a number of exceedingly strong chromospheric flares accompanied by a number of geophysical after-effects: cosmic-ray storms, magnetic storms and atmospheric perturbations. Such flares are extremely dangerous for astronauts in open space and even inside the ship. Unfortunately, we have as yet no way of predicting them.

During World War II many researchers found some radio emission coming from the Sun. Of the radio waves emitted by the Sun we can only receive wavelengths from about 10 metres to several centimetres.

Assuming that within the radio range the Sun is a black-body emitter, then from the intensity of its 1-metre radiowaves its temperature will be hundreds of thousands of degrees. Metre waves are produced by the solar corona, and centimetre waves by the chromosphere. The above-mentioned "temperature" only characterizes the velocity of electrons in these envelopes of the Sun and corresponds to what has been said earlier in the book about the ionization of gases in the corona.

Every so often the solar radio emission increases hundreds of thousands of times. These occurrences are called *flashes*, they accompany large sun-spots, or rather the sudden erruptions of exceedingly hot gases from inside that are called chromospheric flares.

According to Shklovsky, these flashes are due to the fact that the flows of electrically charged particles ejected by the Sun and responsible for polar aurorae on Earth cause in the solar atmosphere some "natural oscillations" of the electrons. These oscillations give rise to sudden radio flashes.

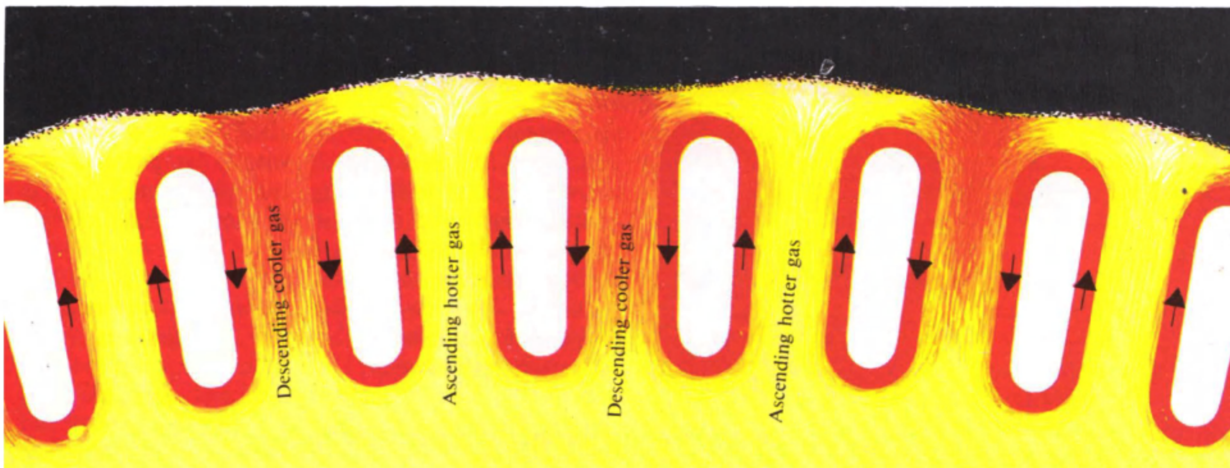
The radio emission of the active Sun is fairly complex in composition, and its theoretical explanation awaits further theoretical and experimental efforts.

Magnetic Phenomena on the Sun

In recent years the theory of the structure of the Sun and the phenomena occurring on it has advanced considerably. In particular, on the basis of laboratory experiments with plasma it has been concluded that the magnetic fields on the Sun play a very important part in the phenomena observed there.

All nuclear reactions occur at the centre of the Sun where the temperatures are sufficiently high – 16 million degrees. The radius of this zone is apparently about 200,000 kilometres. The temperature falls off rapidly – by 20° per kilometre – with distance from the centre. In this region there is a transport of energy by radiation. Before reaching a distance of one-tenth the radius to the photosphere, the temperature drops slowly down, and convection in the form of the ascent of hot gases and descent of cold gases begins to participate in the transport of energy. There is a mixing of matter, which, however, is nonuniform in different directions.

The greater part of the hydrogen atoms in the photosphere are neutral; in the chromosphere, which is essentially a transitional layer, they are ionized; and in the corona the ionization is complete. The photosphere is about 200–300 kilometres thick, that is, about 1/300 of the solar radius. The solar atmosphere, therefore, consists of plasma, a mixture of ions and free electrons. The chromosphere, which is hundreds of thousands of times less dense than the photosphere, thins out into the corona. Being exposed to the emission by the photosphere, with its temperature of 6,000°, the thermometer in the chromosphere would read 5,000° and in the corona still less. The particles of



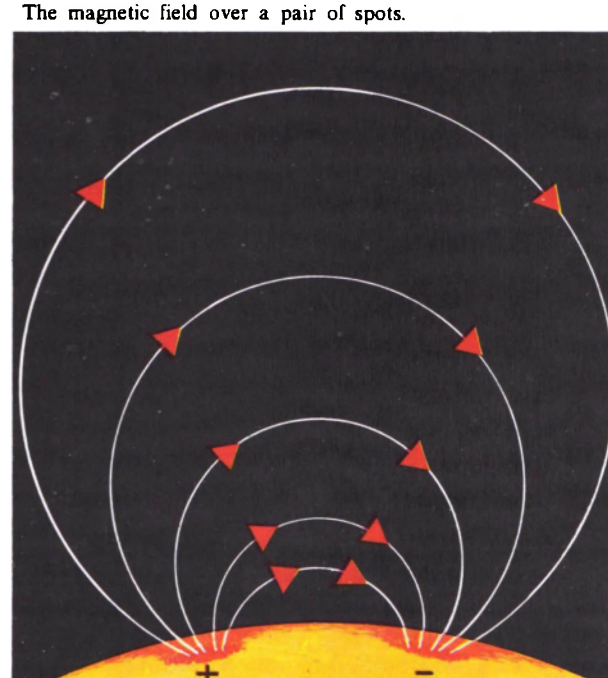
The solar convection zone.

tenuous gas of the chromosphere and corona would encounter the thermometer so infrequently that they could not heat it. However, the velocities of motion of particles in the chromosphere and corona are very high. It is known that the temperature of a gas can be measured by the kinetic energy of its particles. This is the so-called kinetic temperature. In the photosphere the radiation and kinetic temperatures are related to each other, but in the chromosphere and corona they differ markedly – in the chromosphere the kinetic temperature is tens of thousands of degrees, but in the corona it is about a million degrees.

The "heating" of the chromosphere occurs from the energy of the waves propagating in it, waves induced by the motion of granules in the photosphere. In the corona, extending for a distance of up to 10 solar radii, the quantity of atoms per cubic centimetre is 1/100,000,000,000 the quantity of molecules in a cubic centimetre of air at the Earth's surface. At the density of the air, the matter in the corona would only suffice for a layer on the Sun with a thickness of several centimetres. It is here that the principal radio emission of the Sun arises. With the intensity of the corona, a heated body of the same size would radiate at a temperature of a million degrees, but we have seen that such kinetic temperatures also correspond to the bright lines of multiply ionized metals.

A study of the interaction of the magnetic field and plasma has revealed that movement along the lines of force of the magnetic field does not influence the plasma as a whole. But when electrically charged particles move across the field lines (that is, during the flow of a current), an additional magnetic field will

The magnetic field over a pair of spots.



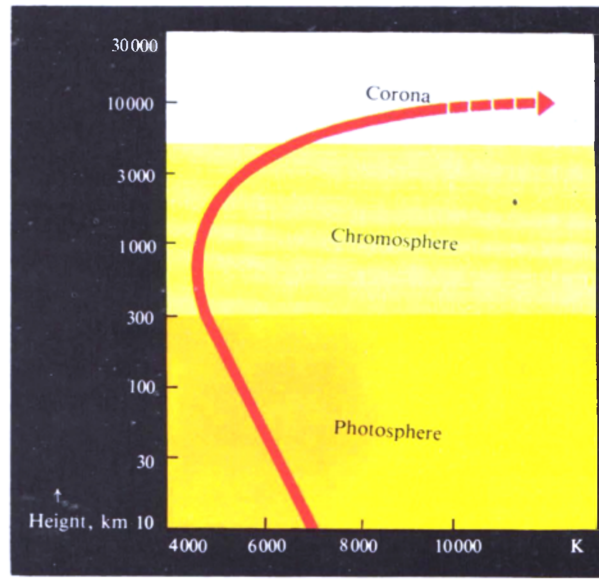
arise. These magnetic fields combined cause the lines of force to curve and elongate following the motion of matter. But the magnetic lines of force tend to straighten out. This creates magnetic pressure, and the field, hindering the plasma from crossing the lines of force, slows it down and can even entrain it, provided the field is strong. If the field is weak, the plasma shifts the lines of force along with it. Thus, in all the cases it can be said that the lines of force are "frozen" in the plasma.

This information, and also regular measurements of the magnetic field strength throughout the Sun's surface, have made it possible to come closer to an explanation of many solar phenomena.

The general magnetic field of the Sun is very weak, but it apparently plays an important role. The coronal streamers, especially in the polar regions, are arranged much like the lines of force emerging from and coming into the poles of a magnetized sphere. The change in field direction in each of the solar hemispheres from one solar cycle to the next is very important as well. The reason for this change is as yet unclear, but stars are known with very powerful magnetic fields in which the polarity of the field also changes periodically.

During the rotation of the Sun the fastest (equatorial) layers entrain the lines of force of the weak general field which are "frozen-in". These lines extend beneath the photosphere and in three years wind around the Sun six times, forming a tight spiral. If in this case the field lines are spaced closer together, this means that the general (and distorted) magnetic field has intensified.

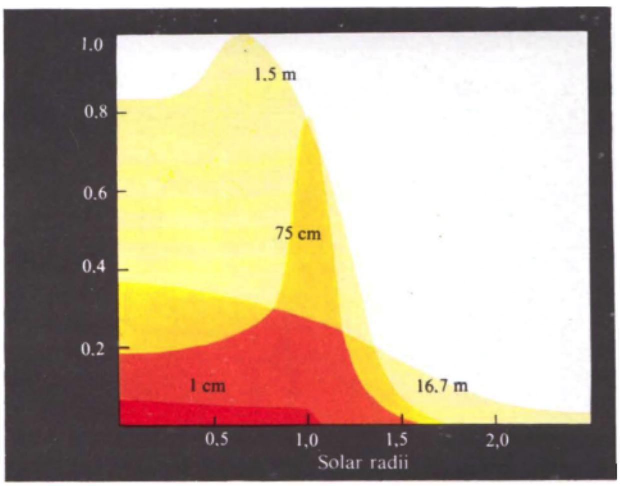
Nearer to the poles the general field lines emerge upward from the photosphere and therefore the field here is not intensified. At the equator itself, where the angular velocity of rotation in a certain zone changes slowly, the field is not intensified either, but at latitudes of $+30°$, where the velocity of rotation changes the most, the field intensification is maximum. Thus, sort of tubes of closely-spaced lines are formed beneath the photosphere. The gas pressure in these tubes adds to the pressure of the magnetic field, perpendicular to its lines. The gas in the tube expands, becomes lighter and can now "surface". Where the tube approaches the surface, the magnetic field becomes more intensive, a facula and then a plage develop. Their hot gases rise above the adjacent parts of the photosphere because the weak magnetic field around them suppresses minor



The variation of temperature with height in the solar atmosphere.

turbulent motions, that tend to slow down the outflow of hot gases. Chromospheric matter above faculae heats up, producing hot flocculi. Finally, the region above the flocculi in the corona becomes brighter, and so develops an active region. When it comes up and breaks through the surface, the tube with closely spaced lines enhances the magnetic field locally, and sun-spots develop. Their temperature is lower because the very strong magnetic field in this region suppresses not only turbulence, but also strong convective movements. The inflow of hot gases from below therefore ceases here, whereas around the spot – in the region of faculae and flocculi – convection is enhanced by the weak magnetic field, thereby increasing the supply of hot gases from below. Clearly, the two places where the curved tube breaks through the surface will have opposite magnetic polarities. When the tube emerges from the photosphere and its lines disperse, the above two spots will break up and disappear. And when the field lines emerge into the tenuous chromosphere and corona, where the gas pressure is less than that of the magnetic field, thc lines diverge and loops and arches form.

Gradually, the active regions, with the magnetic



The variation of radio emission intensity with distance from the Sun's centre (quiet Sun).

tubes producing them, form spots in the eastern part with opposite polarities to those at the beginning of the cycle at this pole. This first causes a neutralization of the former general magnetic field, and then, three years before the end of the 11-year cycle of solar activity, creates a general field of the opposite polarity. Eleven years later the earlier pattern of the general field polarity restores.

Such in broad outline is an apparently correct explanation (due to Babcock) of the 22-year solar periods.

Chromospheric flares on the Sun occur near the neutral points of magnetic fields in active regions, where field strength increases rapidly with increasing distance from these points. Here there is an extremely rapid compression of the magnetic field, together with the plasma into which it is "frozen", and the magnetic field energy in this process is expended as radiation of gas. The plasma is compressed into a thin filament and its temperature increases sharply – to several tens of thousands of degrees. The density of the chromosphere increases here by hundreds of thousands of times in a few minutes.

In addition to the enormous increase in temperature, UV and X-radiation, a chromospheric flare is also accompanied by a so-called radio emission burst, in which radio emission at metre waves jumps tens of millions of times.

The radio source moves from the chromosphere into the corona at a velocity of about 1,000 kilometres per second. It probably arises as a result of an ejection of cosmic rays, generated by the flare, and the bombardment of plasma by these rays, which induces oscillations of plasma giving rise to the radio burst.

The coronal streamers observed are apparently caused by those fluxes of fast, electrically charged particles associated with the magnetic field lines. Both this field and the coronal plasma trap the fluxes of particles, but some of them burst from the solar atmosphere and enter the Earth's atmosphere, where they produce aurorae. It is the change of the pattern of the Sun's magnetic field from its minimum activity to its maximum that governs the changes in the configuration of the corona.

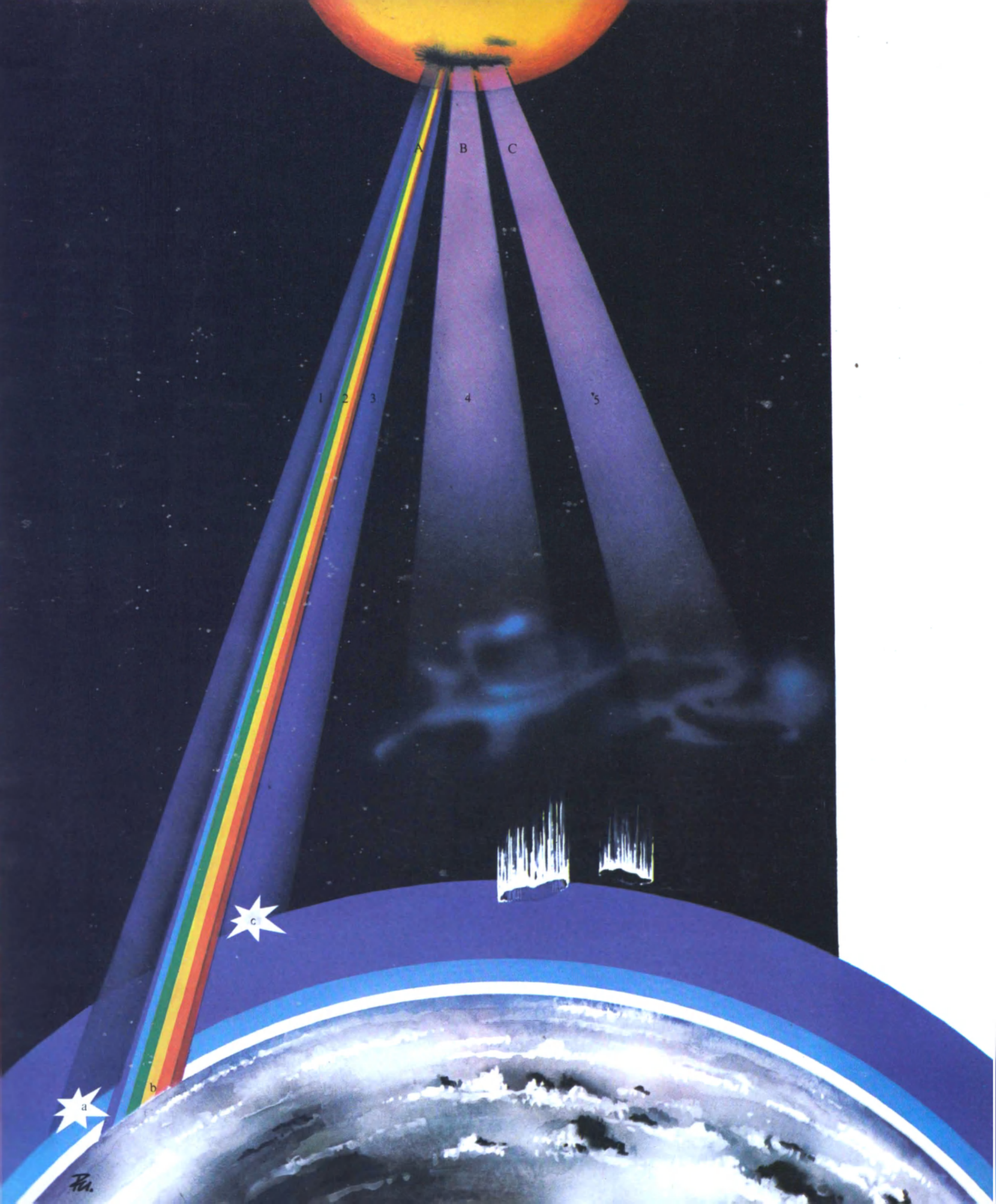
Many prominences, like streamers, are produced by the motion of gas along the lines of force, which results, for example, in surges in archlike paths, whereupon they "slide down" back on the solar surface. Prominences seem to occur for the most part in regions where the magnetic field changes smoothly. The sudden development of the luminescence of prominences at top and then their movement exclusively downward apparently can be attributed to processes similar to those responsible for chromospheric flares, but in this case the processes are less pronounced. The compression of the magnetic field leads to the compression of relatively cool gas, to an increase in its density, and to luminescence.

Such in its essentials is the present-day, largely magnetohydrodynamic, theory of solar phenomena.

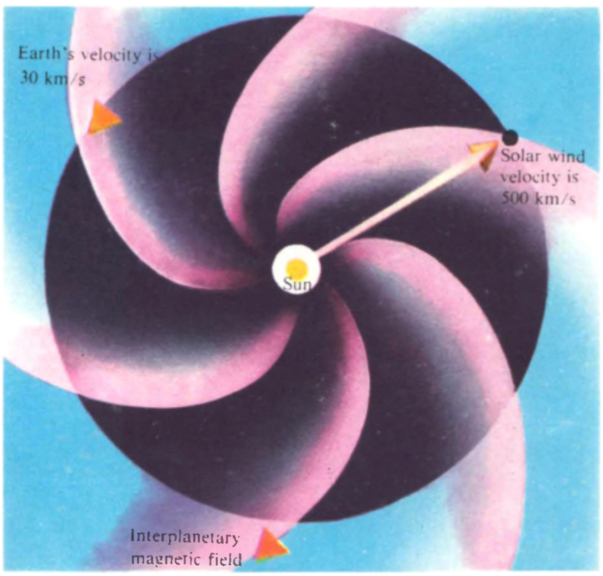
The Solar Wind and Polar Aurorae

The Norwegian geophysicists Birkeland and Störmer have long concluded that the Sun emits flows of electrically charged particles. They argued that these particles, on hitting the Earth's atmosphere, could produce perturbations of the terrestrial magnetic field and polar aurorae. During sun-spot maximum periods

Terrestrial implications of a large solar flare:
A – electromagnetic radiation (8 min); *B* – cosmic rays (less than 1 hour); *C* – corpuscular radiation (from 20 to 40 hours). At first come: *1* – ultraviolet radiation, X-rays, and gamma-rays; *2* – light; *3* – radio waves. Then come: *4* – protons (and other nuclei); *5* – ions and electrons.
Produced are the following effects: *a* – sudden ionospheric disturbances in the so-called D-layer (Mögel-Dellinger effect); *b* – H_α-flares; *c* – flares in the high atmosphere.



A
B
C
1
2
3
4
5
a
b
c



Schematic diagram of the interplanetary magnetic field (the black point is the Earth).

aurorae and magnetic storms on Earth become more frequent and intensive. The latter manifest themselves in field fluctuations and vibrations of the magnetic pointer of a compass. A German astrophysicist, Biermann, showed in the 1950s that a number of phenomena in cometary shapes, specifically notable accelerations of gases in type I cometary tails, must be the result of the interaction of the plasma of a cometary tail with solar corpuscular flows and the magnetic field associated with these latter. The theory is developing further and now it explains the many phenomena that could not be accounted for when it was believed that comets are chiefly affected by the pressure of sunlight. Since comets continuously wander around the solar system, it can be concluded that the Sun's corpuscular radiation is emitted continuously and in all direction, thus filling the solar system.

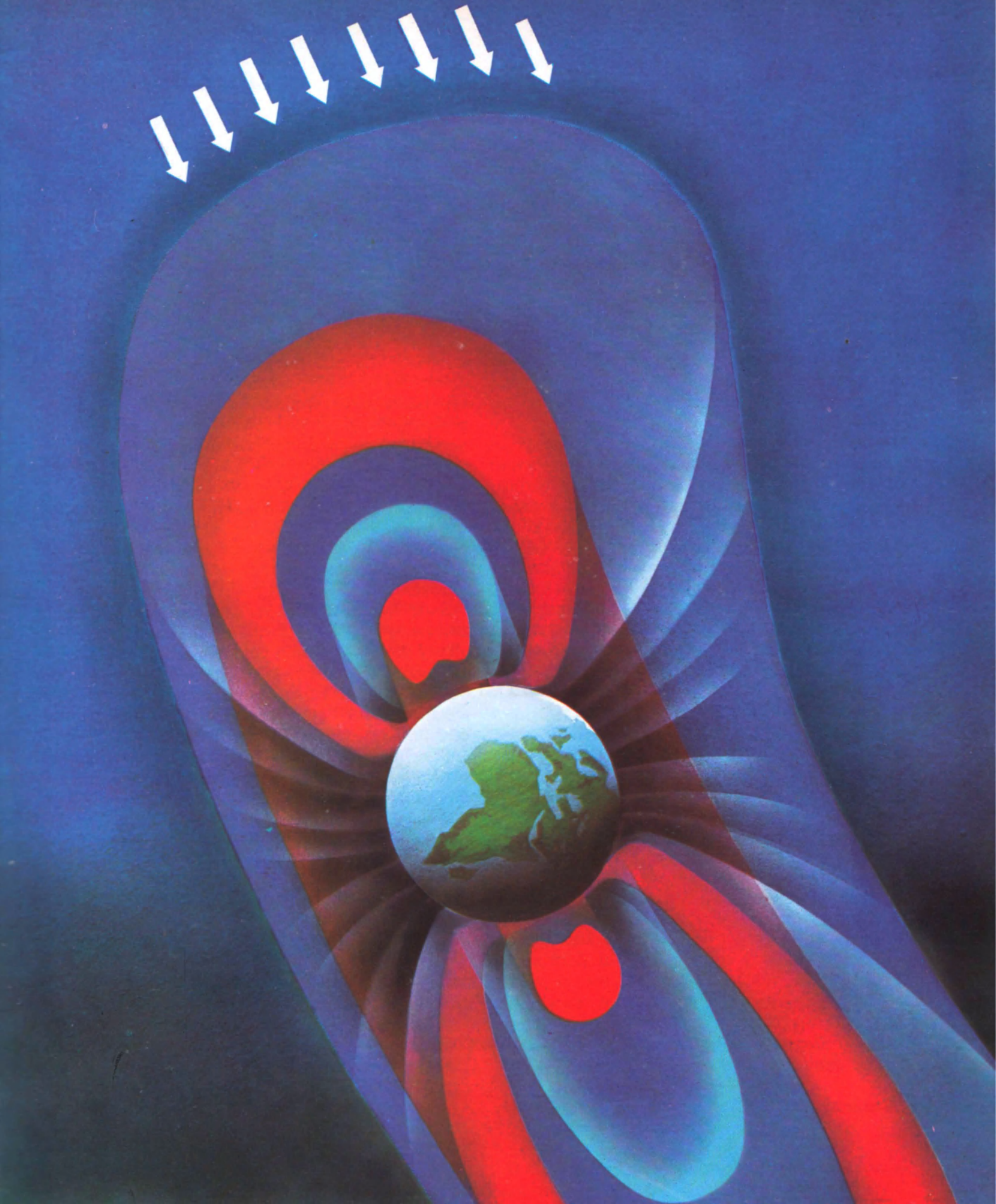
A team of Soviet authors associated the corpuscular flows with the coronal streamers. Treating the corona as a dynamic formation, they decided that it expands continuously.

It was later concluded that the solar corona, having a temperature of a million degrees, should be extended by a large thermal flow it produces to the Earth's orbit, where its temperature drops down to 200,000 °C. The corona density here was estimated at 100-1,000 hydrogen atoms per cubic centimetre, which also followed from the explanation of the polarization of zodiacal light. Astronomers now used to say: "We live within the solar corona." Even at a distance of 10^7 kilometres from the Sun the expanding corona must move at several hundreds of kilometres per second, and so the expanding solar corona was identified with the Sun's corpuscular radiation, hence the name "solar wind".

Also, experimental evidence began to be gleaned: beginning in 1959 the planetary probes have been recording the solar corpuscular flows in interplanetary space at different distances from the Earth. So it has been found that at the Earth-to-Sun distance the solar wind "blows" constantly with a velocity of about 400 kilometres per second and that the number of particles in one cubic centimetre varies somewhat. Normally, 10^8-10^9 particles cross one square centimetre per second after large flares on the Sun. The flow carries a magnetic field, the particles moving along spirals, not radii. But the solar wind is by no means constant. It is characterized by gaseous turbulence and field deformation. At present the solar wind is being studied extensively using all the means available, since it is responsible for both cometary processes and many geophysical phenomena in the terrestrial magnitosphere and atmosphere.

Polar aurorae are most commonly, almost every day, observed in ring zones with a radius of 23° and the centre near the magnetic poles of the Earth. The strongest and highest aurorae are at times observed not only in the high and medium altitudes, but also in the tropics, when they accompany magnetic storms noted throughout the world. The types of aurorae, those electric stratospheric luminescences, vary widely. Although produced by energetic particles related to solar activity, they vary in the causes of different kinds of luminescence. Their altitude is determined from their parallax. The commonest height is 95-120 kilometres, but sometimes somewhat lower. Rarely enough, aurorae are observable at altitudes up to 1,000 kilometres. Knowledge of the height and air density at different heights enables the velocity and energy of the particles in the stratosphere to be worked out. For

Van Allen belts.



example, protons with an energy of 100 kiloelectron-volts and electrons with an energy that is even 1/10 of that figure are able to penetrate down to a height of 100 kilometres.

Auroral spectra show bright lines of atomic oxygen and nitrogen and bands of molecular oxygen and nitrogen, both neutral and ionized. Among them there are also forbidden lines discussed earlier in the book. They are caused by the thin stratosphere which at 100 kilometres is 1/1,000,000 as thin as the air at sea level.

Individual lines are mostly caused by the difference in the chemical composition of the air at various altitudes and in the energy of the incoming corpuscules. That is why polar aurorae display a stunning play of colours and their gleam on the white of the arctic and antarctic snow wastes. In addition, the auroral spectrum has the red line of atomic hydrogen, which is due to the solar protons.

Our understanding of the nature of aurorae gained enormously from the extensive systematic studies conducted under the umbrella of the International Geophysical Year, and also from orbit-borne observations.

The terrestrial magnetic field resembles the field of a magnetized iron ball with the lines of force coming out of one magnetic field and into the other. The arcs of polar aurorae appear to be extended along geomagnetic parallels and their rays along the geomagnetic lines of force.

Amorphous luminescence near the geomagnetic poles is produced by highly energetic electrons coming directly from the Sun and deflected by the terrestrial magnetic field. The aurorae within the region of their most frequent occurrence (polar circles) are excited by electrons with an energy of no less than 10 kiloelectron-volts, which cannot come directly from the Sun. They acquired their high velocity while travelling in the Earth's magnetic field, although they do not belong to the terrestrial radiation belts. Here also occur aurorae caused by 1.5-2 kiloelectron-volt protons belonging to the solar wind, which first stream around the geomagnetic field and then rush into the stratosphere. Responsible for red aurorae at an altitude of about 350 kilometres that occur closer to the Earth's equator are low-energy protons of the solar wind. How the particles of the solar wind penetrate into the stratosphere and what happens to them in the process are the questions that await their answer.

Solar wind produces magnetic storms and polar aurorae, perturbances in the Earth's ionosphere, and affects short-wave radio communication and maybe, according to Chizhevsky, it exerts a notable influence on living things. This makes the studies of the solar wind and related phenomena especially important in connection with manned space flights.

Since the solar wind is related to the active regions on the Sun, existing for a long period of time, and the Sun's rotation speed and the speed of the corpuscular flows are known, we can to a certain degree predict magnetic storms and intense polar aurorae.

2. The Stars Are Distant Suns

Comprehending the Incomprehensible

The "collective author" Koz'ma Prutkov has presented the following anecdote about Descartes:

"Once when the night covered the heavens with its invisible shroud, the famed philosopher Descartes, sitting on a step of his home, was approached by a passing person in the street who pointed to the dark horizon and said: 'Tell me, wise man, how many stars are in the heavens?' 'Idiot!' he responded, 'no one can comprehend the incomprehensible'".

The meaning of these words is that the stars visible in the sky have no number, and if we refer to the stars visible with the naked eye, they all have been counted long ago. This problem is not incomprehensible. It is entirely possible to comprehend the number of stars.

Look at the starry sky, search in it constellations using a star map, and you will soon see how easy it is to take bearings in the sky and be able to count all the stars visible with the naked eye. There are only about 6,000 of them, but immediately over the horizon it is only possible to see about 3,000. If we say "about" it is only because the acuity of vision and the transparency of the air can be different. The lists and maps give not only all these stars, but a great many fainter ones.

The fainter the stars, the more of them are there in the skies and even their simple counting becomes increasingly difficult.

All the stars brighter than the 11th magnitude have been counted and entered in catalogues, so to speak, one by one. The number of fainter stars is also known, although not so precisely, but this is not so important. We deal with them like a forester with trees in the woods, not counting each tree when he computes the volume of timber. In small standard areas of a determined size the forester counts the number of trees and then multiplies this by the number of such areas in the expanse occupied by the forest. We proceed in the same way with the stars.

As a result, the counts of stars brighter than a particular stellar magnitude can be represented by the table that follows.

Thus, we hold to strict account about a million stars, and in all we are able to observe about 2,000,000,000. The numbers are breath-taking, but they can be comprehended.

Maximum stellar magnitude	Number of stars
6.0	4,850
7.0	14,300
8.0	41,000
9.0	117,000
10.0	324,000
11.0	870,000
13.0	5,700,000
15.0	32,000,000
17.0	150,000,000
19.0	560,000,000
21.0	2,000,000,000

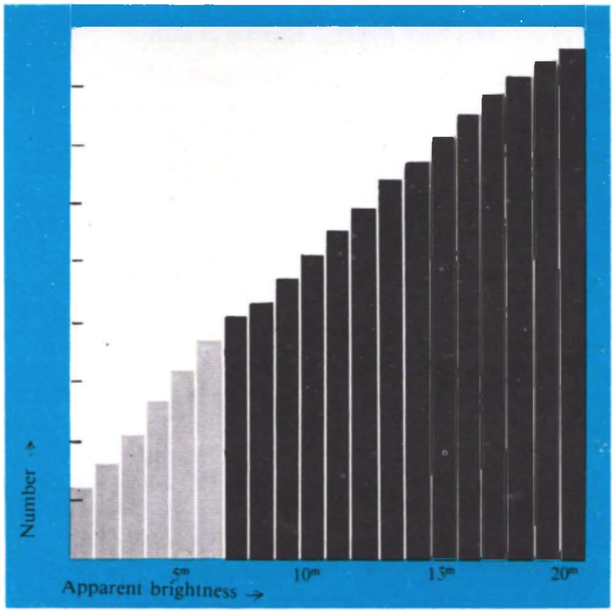
Stellar Luminosity

Somewhere at sea in the darkness of the night twinkles a little light and if some old seaman does not explain to you what it is you frequently will not know whether it is a lantern on the bow of a passing gig or a powerful light on a distant lighthouse. This is the position we are in on a dark night when looking at the scintillating stars. Their apparent brightness is dependent both on the true intensity of their light, called *luminosity*, and on their distance from us. Only a knowledge of the distance to a star makes it possible to compute its luminosity in comparison with the Sun. For example, the luminosity of a star ten times less bright than the Sun is expressed by the number 0.1.

The true intensity of a star's light can also be expressed differently by computing what stellar magnitude it would appear to us if it were situated away from the Earth at the standard distance of 32.6 light-years. A tenth of this distance (3.26 light-years) is used by professional astronomers as a unit for interstellar distances and is called the *parsec*. It is so called because from this distance the angle at which it is possible to see the radius of the Earth's orbit, perpendicular to the line of sight (this angle is called the *annual parallax*), is exactly one *second*. The parsec is 206,265 times greater than the distance from the Earth to the Sun (i.e. an astronomical unit), so that

1 parsec = 3.26 light-years = 206,265 astronomical units = $3.083 \cdot 10^{13}$ kilometres.

At the standard distance of 10 parsecs, or 32.6 light-years, the Sun would appear to us as a star of the 5th



Counts of stars (log scale) of various apparent brightnesses.

stellar magnitude, that is, it would not be particularly easily visible with the naked eye even on a moonless night. The stellar magnitude of a celestial body at this standard distance is called the *absolute stellar magnitude*.

The brightness of stars, like that of any light source, changes inversely with the square of distance. This law makes it possible to compute the absolute stellar magnitudes or luminosities of stars knowing the distance to them.

Suppose, for example, that a star of the 5th apparent magnitude is situated at a distance of 40 parsecs. Then at a standard distance of 10 parsecs it would be four times closer to us and its apparent brightness would increase 4^2, i.e. 16 times. But 16 is almost precisely 2.5^3, i.e. the stellar magnitude of the star would become 3 units less. Instead of the 5th, it would be of the 2nd stellar magnitude, i.e. 3 units brighter than the Sun ($M_\odot = +5$). Consequently, its luminosity is

$$2.5 \times 2.5 \times 2.5 = 2.5^3 = 16\,.$$

Thus, this 5th magnitude star in reality is 16 times brighter than the Sun. Absolute magnitude M of a star is readily computable from its apparent stellar magnitude m and the distance D in light-years by using the formula

$$M = m + 7.5 - 5 \log D\,.$$

When the distances to many stars became known, we could compute their luminosities, i.e. we could arrange them in a row and compare them with one another under identical conditions. It must be admitted that the results were stunning, because it had earlier been assumed that all stars were "similar to our Sun". The luminosities of stars were found to be surprisingly different. We will only cite extreme examples of luminosity in the world of stars.

One of the faintest is star No. 359 in the Wolf catalogue. It is 50,000 times fainter than the Sun and its absolute magnitude is $+16.6$.

At the other end of the star list is S Dor, visible only in the countries of the southern hemisphere as a faint star of the 8th magnitude. It is 1,000,000 times brighter than the Sun and its absolute magnitude is -10.6. If the brightness of an ordinary candle were assumed equal to the brightness of the Sun, by comparison S Dor would be a powerful searchlight and the star 359 in the Wolf catalogue is fainter than the tiniest glow-worm!

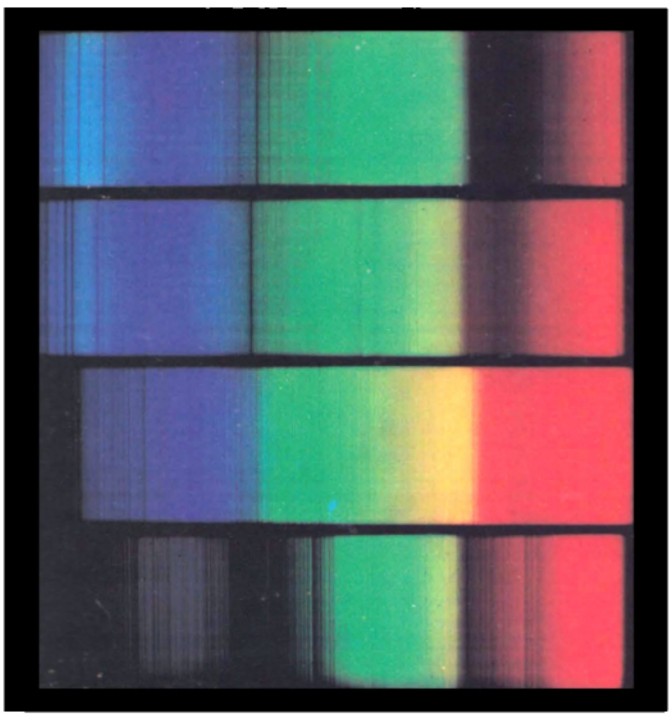
Thus, stars are distant suns, but their luminosity can be completely different from that of our Sun. If we were to exchange our Sun for another we would have to use caution. The light of one would blind us and with the light from another we would lose our way, like in twilight.

Spectra – Stellar Credentials

Stellar spectra are passports, with a description of all stellar features and all their physical properties. We need only be able to interpret these passports. In the future we will be able to learn much more from them, but even now we can learn a thing or two from them.

From the spectrum of a star we can work out its luminosity (and hence its distance), temperature, dimensions, the chemical composition of its atmosphere (both qualitative and quantitative), the velocity of its motion in space, the velocity of its axial rotation, and even whether there is another invisible star nearby that accompanies it in revolution around a common centre of mass.

If spectra are the passports of stars, it is natural for us to attempt to classify stars, i.e. to obtain the



Stellar spectra (from above); Spica (B 1), Procyon (F 5), Arcturus (K 1), Betelgeuse (M 2).

spectrum for each of them. We have already obtained the spectra of hundreds of thousands of stars, but unlike the luminosities, they appeared to be far less varied, making it possible to break them down into a small number of spectral classes.

Each spectral class corresponds to a particular combination of physical properties of a star, so that when we say that a star belongs to spectral class A, we immediately give some idea of its physical nature. For example, an ordinary star of spectral class A will be dozens of times brighter than the Sun; of a white colour; with a surface temperature of about 10,000°, and several times greater in diameter than the Sun. The most conspicuous lines in its spectrum will be the dark lines of hydrogen. This forces us to pay particular attention to the spectral classification of the stars.

We note first that the stars are of different colours, of all shades from red through yellow to white and bluish. A particular colour of a star corresponds to a particular spectral form. They are all similar to the solar spectrum in that dark lines are visible against the background of a continuous spectrum. But in addition to spectra that contain the same lines as are present in the solar spectrum there are spectra with completely different lines.

On the basis of a proposal of the Harvard Observatory the spectra are classified by the intensity of spectral lines, and so they are arranged in decreasing order of temperature of stars, or rather in the order of their change in colour from bluish and white to red, the temperature of stars changing together with their colour.

The table below gives a comparison of the principal characteristics of stars, associated with their spectrum.

Spectral Classes of Stars

Spectral class	Colour	Temperature, C	The most intensive lines	Typical bright stars
O5	Bluish	30,000	Ionized helium	—
B0	White	20,000	Helium	β Cruc
A0	White	10,000	Hydrogen	Sirius, Vega
F0	Yellowish	8,000	Ionized metals	Canopus,
G0	Yellow	6,000	Neutral metals	Capella, Sun
K0	Orange	4.500	Weak bands of titanium oxide	Arcturus
M0	Red	3.000	Intense bands of titanium oxide predominate	Antares

The stars in the upper part of the table are called early (purely arbitrarily) and the others late. Spectra intermediate between these are denoted thus: a spectrum a little later than A0 is denoted A1, still a little later – A2, intermediate between A0 and F0-A5, slightly earlier than F0-A9, and so on. No spectra earlier than O5 have been found.

We note that among cold red stars, there are two other varieties in addition to class M. In the spectra of certain stars the bands of molecular absorption of titanium oxide are replaced by bands of carbon monoxide and cyanogen (in spectra denoted by the letters R and N), and still others are characterized by bands of zirconium oxide (class S).

The spectrum of each star has a great many lines of different chemical elements, and the table lists only the most characteristic ones.

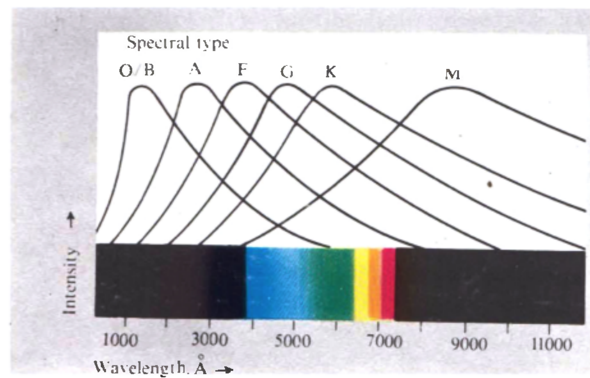
What Do Stars Consist of and Why Are Their Spectra Different?

A difference in spectra – the stellar passports – was known long ago and was correctly explained in the 1930s using the theory of ionization of gases. It was found that this difference is dependent for the most part not on the chemical composition of their atmosphere, since the atmospheres of all stars are almost identical, but on the difference in their temperatures. So at the relatively low temperature of K and M stars or sun-spots can exist stable chemical compounds such as titanium oxide. In a hotter star titanium oxide breaks down into its components – titanium and oxygen. In its atmosphere the atoms of metals, easily excited and willingly absorbing the light of the photosphere, will be the principal absorbers and will stand out most clearly in the spectrum. In a still hotter star, the atoms of metals are ionized and yield different lines.

In a yet hotter star, the atoms of metals lose more than one electron and their spectral lines move into the invisible ultraviolet part of the spectrum, leaving room for the hydrogen lines. Hydrogen atoms are excited in greater quantity and absorb the light of the photosphere (and thereby produce dark lines in the spectrum) and do so more intensively at the higher temperature of the star.

The chemical composition of the atmospheres of stars and the Sun for the most part is almost identical and close to that of the Earth's crust, both quantitatively and qualitatively, with the exception that there are no appreciable quantities of hydrogen and helium in the Earth's atmosphere. Abundances of atoms in stellar atmospheres can now be determined from the intensities of the dark lines they produce – according to the theory of atomic spectra and laboratory experiments to determine the absorptivity of various gases and vapours. But stellar atmospheres vary somewhat in chemical composition as exemplified, for example, by the difference in the spectra of M and N stars and in hot Wolf-Rayet stars.

The third table in this chapter gives the logarithms of the mean quantity of atoms in a column of the atmosphere with a cross section of 1 square centimetre for stars and the Sun in comparison with the same, but relative, data for the Earth and



The continuous distribution of intensities of solar spectra.

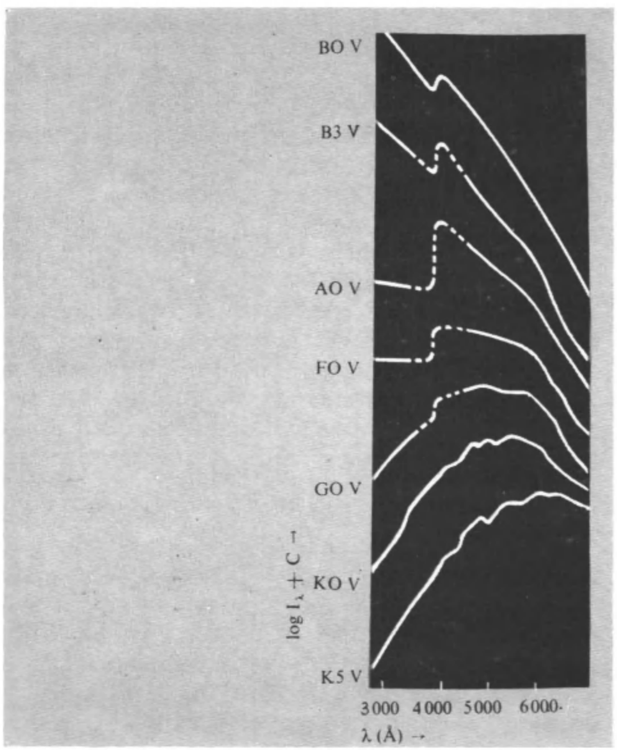
meteorites. (It should be recalled that if logarithms differ by 2, the numbers differ by hundred times, etc.)

Comparative Abundances of Chemical Elements in the Atmospheres of Various Suns, Including Our Own, in the Earth's Crust and in Stony Meteorites

	Stars	Sun	Earth's crust	Stony meteorites
Hydrogen	11.4	11.5	8.3	6.9
Helium	10.2	10.2	0	0
Carbon	6.4	7.4	6.3	6.1
Oxygen	8.0	9.0	8.5	8.4
Sodium	7.1	7.2	7.3	6.4
Magnesium	7.5	7.8	7.2	7.7
Aluminium	6.9	6.4	7.8	6.8
Silicon	7.5	7.3	8.2	7.8
Potassium	5.3	6.8	7.1	5.6
Calcium	6.7	6.7	7.2	6.5
Titanium	6.0	5.2	6.4	5.3
Vanadium	4.9	5.0	5.2	?
Chromium	5.8	5.7	5.4	5.8
Manganese	6.5	5.9	5.6	5.6
Iron	6.7	7.2	7.2	7.6

All these data are not entirely precise, but we see that the atmospheres of star-suns not only consist of the chemical elements as the Earth's crust, but also that the relative content of each of them in the Earth and in the stars is very similar, except for the fact that stars and the Sun are far richer in hydrogen and helium.

It is interesting to note in this connection that the philosopher Auguste Comte – a positivist and ide-



Variation of intensity with wavelength for various types of stars.

alist – on the eve of the discovery of spectral analysis asserted that man would never know the chemical composition of stars. Science has refuted this and many other predictions, including the prediction that it would be impossible to observe the solar corona at times other than during total eclipses. It is impossible to set limits to the human quest for knowledge!

Not all the known chemical elements have been discovered in the atmospheres of stars and the reason for this is the same as we have noted for the Sun. The atmospheres of stars and the Sun differ from the Earth's primarily in that they are richer in hydrogen and helium. It has been found that the interiors of stars, at least most of them, also largely consist of hydrogen.

The chemical composition of some stars deviates from the average. For example, there are stars that are somewhat richer in neon or strontium. Shain at Simeiz discovered that certain cool stars feature anomalously great abundances of a special form of carbon, a so-called heavy isotope of carbon. Although possessing the same chemical properties as ordinary atoms of carbon, the atoms of this isotope are 1/12 heavier. This shows that in different worlds there is no absolute identity of conditions, as should be the case in an infinite and infinitely varied universe.

Of great interest was the discovery in 1965 of several "infrared stars". Although seen visually through the telescope, they have most of their radiation concentrated in the far infrared range of spectrum beginning with 9,500 angströms. They somewhat resemble in spectrum the very red stars, but represent some class of their own of objects with very low temperature (1,000 °C, or maybe even 700 °C). This leads us to assume that some quite dark stars are in existence, but they must be rather few. The spectrum of infrared stars has absorption lines of titanium and vanadium oxides, and besides absorption lines... of water vapour. Yes, water on stars, where it would be the last substance to be expected. But in the outer atmosphere of so cool stars there may be some hot water vapour. You should remember, of course, that these so-called "cool" stars consist of tenuous hot *gas* – they are not molten either, and the more so, not solid.

The comparative uniformity of chemical composition of known celestial bodies will possibly disappoint some people. However, this fact, confirming the material unity of the cosmos, is undoubtedly of great significance. This unity enables us to apply to the stellar universe the laws of nature that we have learned experimentally within the modest limits of our Earth. These facts go to confirm the correctness of the viewpoints of dialectic materialism.

Thermometers for Stars

We have already stated that the temperature of celestial individuals can be determined even without inserting a thermometer. The methods for measuring the temperature of stars are numerous and they mutually check one another. The results obtained by the various methods agree quite well and lead to the conclusions shown in the table of stellar spectra. We can measure the temperature of a star from

enormous distance by intercepting the heat it emits, or we can compute the temperature from the dimensions and luminosities or the relative intensity of the different spectral lines using the theory relating the degree of ionization of a gas to its temperature, checked experimentally. In addition, the temperature can be determined from the energy distribution in continuous spectra and even from the colour of the star, which is dependent on this distribution.

It must be remembered that since the spectrum and light of stars we observe are produced by their atmospheres, both the temperatures and chemical composition we determine only apply to the atmospheres. The temperature within stars is determined by complex theoretical computations; these temperatures attain many millions of degrees.

Stellar Spectrum – Distance Indicator

Knowing the luminosity and apparent brightness of a star, we can readily work out its distance. So using the absolute (M) and apparent (m) stellar magnitudes we obtain the distance in light years D from

$$\log D = \frac{m - M + 7.5}{5}.$$

Stellar spectra were found to be adequate indicators of luminosity, and hence of distance, since the apparent brightness m, which is required for the comparison, is readily obtainable.

And if we could, using some other techniques, determine the distances of a group of stars, we could also work out their luminosities and compare the results with the spectra of the stars.

But at this point it would be sufficient to point out that conventional white stars of some spectral subclass, say A0, A1, A2, etc., have a definite luminosity, generally determined using suitable diagrams.

It is thus sufficient to determine accurately the spectral subclass of a conventional white star to get some indication of its luminosity, and hence distance. (There are A stars of other luminosity, but their spectra are a bit different. Such stars occur fairly rarely.) Why this is so is another story, but this does not prevent us from making use of the fact itself.

Things are far more complex with yellow and red stars, although quite definite as well. Yellow, and especially red, stars of one spectral class are sharply divided into two groups. Some are called giants and have an exceedingly high luminosity, the others are called dwarfs and have much lower luminosity. There are no stars of intermediate luminosity and the luminosity of both dwarfs and giants of the same spectral subclass is quite definite.

But fortunately there is some difference in the spectra of yellow and red dwarfs and giants belonging to the same spectral class. The same dark lines in the spectra of giants are narrower and sharper than in the spectra of dwarfs. This helps us to distinguish them from one another.

Moreover, the relative intensity of *certain pairs of lines* is clearly dependent on the luminosity of a star. The spectra of dwarfs and giants are not identical. For example, the spectra of orange stars 61 Cyg and Aldebaran in general are similar and they belong to the same spectral class K5. But among the numerous identical lines in their spectra it can be noted that the calcium line with a wavelength of 4454 Å in the spectrum of the dwarf 61 Cyg is stronger than the 4215 Å line of ionized strontium, but with the giant Aldebaran the opposite is true. A certain amount of experience is needed to distinguish the spectra of giants and dwarfs from one another. It is possible to establish a dependence between the relative intensity of *pairs of lines* (similar to those mentioned above) and the luminosity of a star and make use of it in later work. Then, having photographed the spectrum of a star of unknown luminosity located at an unknown distance, we can easily establish both its distance and luminosity.

It is remarkable that the accuracy of determination of distances to stars by this method is about 20 per cent, regardless of whether the star is nearby or distant. This is the same as if we could measure with the same relative accuracy the length of a table and the distance from Moscow to Vladivostok. It may be that you think an accuracy of 20 per cent is not sufficiently good for measurement of the distance to stars. We must agree. In most cases, however, it is impossible to determine the distance to a star by any other method. But in order to determine it from their spectra it was necessary to measure many distances by more direct means, which will be discussed later in the book.

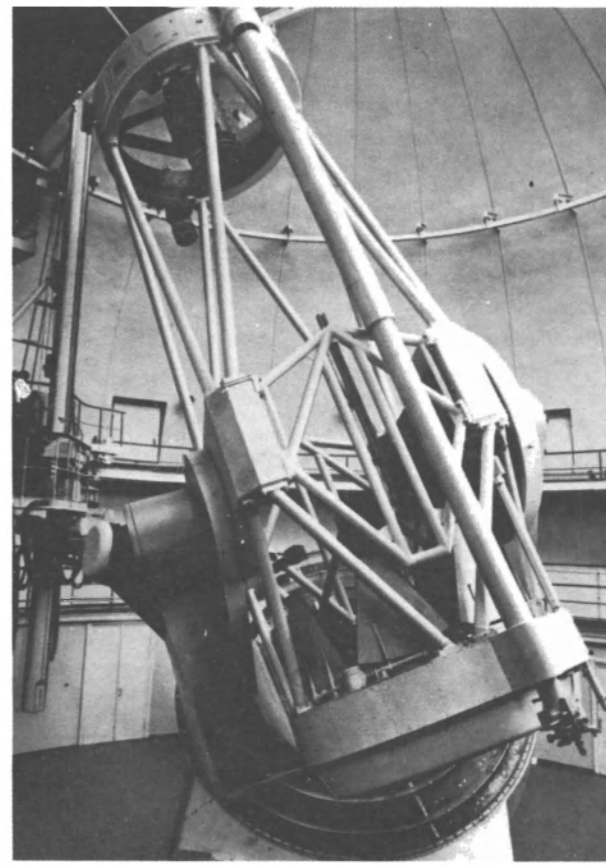
With respect to the physical cause of the secondary differences in the spectra of dwarfs and giants, we find that the atmospheres of giants are more extensive and tenuous, which influences greatly the intensity of certain lines. The remaining lines, while sensitive to changes in gas temperature, are insensitive to changes in pressure.

Sounding Rod in the Depths of Space

Beyond the confines of the solar system it is necessary to make such a big jump in distances to the stars that it could only be done a century ago, far later than all doubt had disappeared as to a similarity between the Sun and the stars. The sounding rod – used for measuring depths at sea – was repeatedly thrust into the direction of various stars but for a long time could not reach a single one of them, could not reach "bottom". This, of course, is only a figure of speech, because as in the case of determination of the temperatures of celestial bodies, the possibility of direct measurements of distances was excluded. As we shall now see, they can only be determined by indirect means, by computations based on measurements of other values. This method, suggested by Copernicus, involves the measurement of angles, but the instruments and techniques making the necessary accuracy possible were not devised until the second half of the 19th century.

As in the determination of distance to any inaccessible object, the method consists in measuring the difference in directions from which a star is visible from two ends of a base of known length. The distance corresponding to this difference in directions can be computed by trigonometry. In this case the Earth's diameter was too small as a base and for the majority of stars, considering the present-day accuracy of angle measurement, even the diameter of the Earth's orbit was inadequate. Nonetheless, it was this which Copernicus recommended to use as a base and this was done by scientists of later generations.

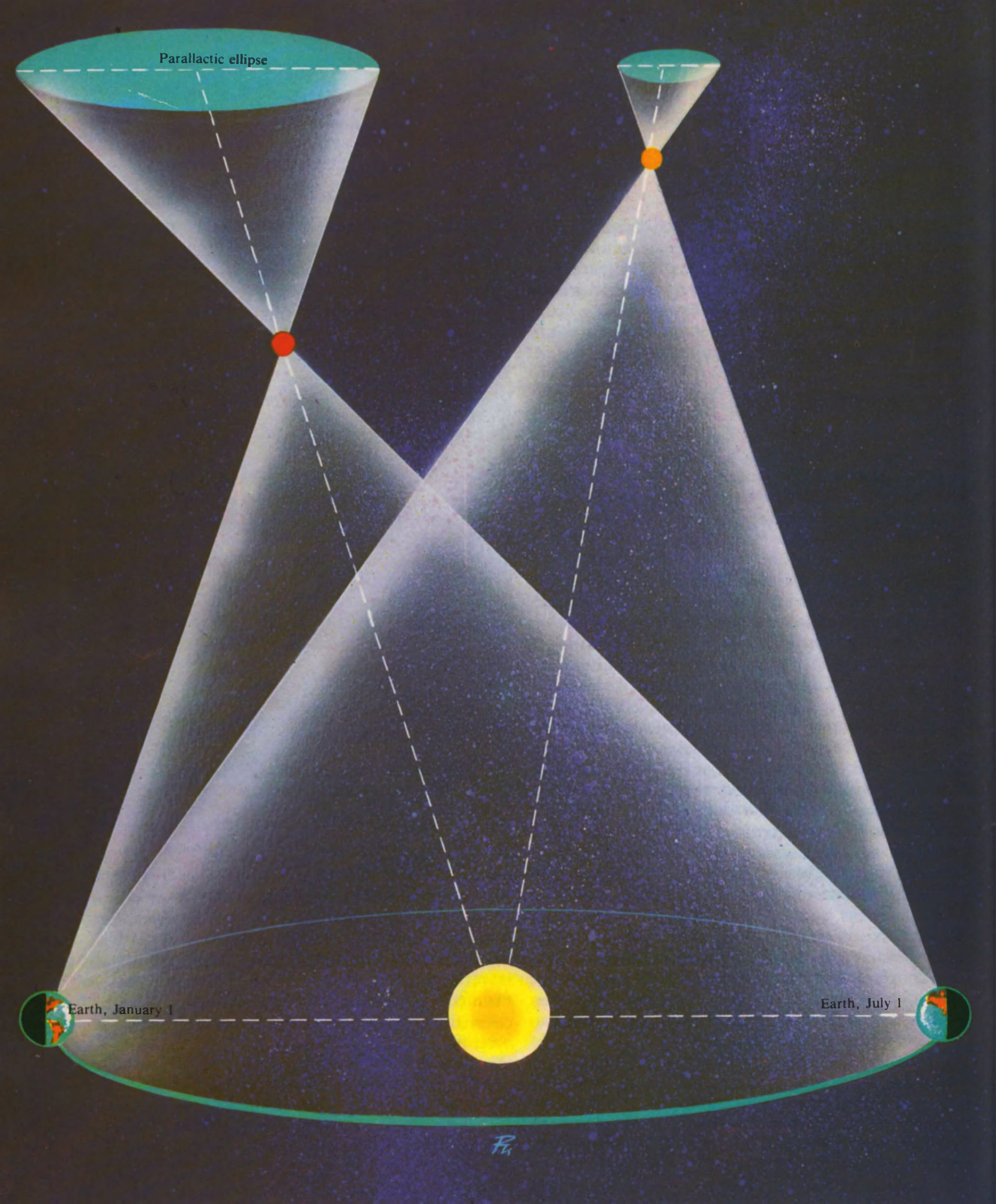
Only 140 years ago were the astronomers Struve of Russia, Bessel of Germany and Henderson of South Africa able to make quite precise measurements and for the first time established the distances to certain stars. The feeling experienced at that time by contemporaries was similar to the joy of seamen who after a long voyage of unsuccessful sounding finally reached bottom.



The 2.6-metre telescope of the Crimean Observatory.

The classical method of determining star distances involves the precise determination of direction to them (i. e. determination of their coordinates in the celestial sphere) from the two extremes of the Earth's orbit at times separated by a half-year, since during this time the Earth has carried the observer from one extreme of its orbit to the other.

The apparent displacement of the star, caused by a change of the position of an observer in space, is extremely small, scarcely perceptible. It is measured preferably on a photograph, by taking on the same plate two pictures of a given star and its neighbour, one after the other in a half-year. The majority of


Parallactic ellipse
Earth, January 1
Earth, July 1

stars are so far away that their displacement in the sky in this case is completely undetectable, but a sufficiently close star will be displaced appreciably relative to it. This displacement of the star is measured with an accuracy to 0″.01. Up until now it has not been possible to achieve a greater accuracy, but it is already far greater than the accuracy attained half a century ago.

The apparent displacement of a star is twice as large as the angle known as the annual parallax. (In discussing the annual parallax we can speak of the radius of the Earth's orbit because we can here ignore the slight ellipticity of the latter).

Our closest neighbour is the first magnitude star α Centauri, although one close to it, invisible with the naked eye, is even 1 per cent closer. The parallax of these stars is maximal and is 3/4″; it has been measured with an accuracy of about 1 per cent, since the accuracy of angle measurements attains 0″.01.

An angle of about 0″.01! This is equivalent to viewing from Birmingham a penny piece set on edge in Trafalgar Square in London. Such is the accuracy of astronomical measurements.

From the parallax it is easy to determine the distance. We obtain the distance to a star in radii of the Earth's orbit if we divide the number 206,265 by the parallax, expressed in seconds of arc. In order to express it in kilometres the number derived must be multiplied by 150,000,000.

We already know that great distances are more conveniently expressed in light-years or in parsecs. The star α Cen and its neighbour, which is called Proxima Centauri, or the "closest", are 270,000 times farther away than the Sun, i.e. 4 light-years. An express train travelling at 100 kilometres per hour would take 40,000,000 years to cover the distance. Try to console yourself with this fact if you ever become tired of a long train ride.

An accuracy of parallax measurement of 0″.01 does not permit measurement of parallaxes less than this value, so that the method described is not applicable to stars farther away than 300-350 light-years.

It is easy to see that if at a parallax of 0″.01 the error of 0″.01 is 10 per cent, at a measured parallax of 0″.01 we should assume that the true parallax can be both 0″.02 and 0″.00. In the first case the true distance (corresponding to 0″.02) is about 150 light-years, in the second case it is infinite. With these small parallaxes an error of 20 per cent in the determination of distances from the spectrum makes them more reliable than the classical method of determination of distances, on which, however, they themselves are based. The classical method, described here, has been used during 150 years to determine the distance to 6,000 stars (about 10,000 determinations) but from spectroscopic parallaxes the distances of about 28,000 objects have been determined during 30 years.

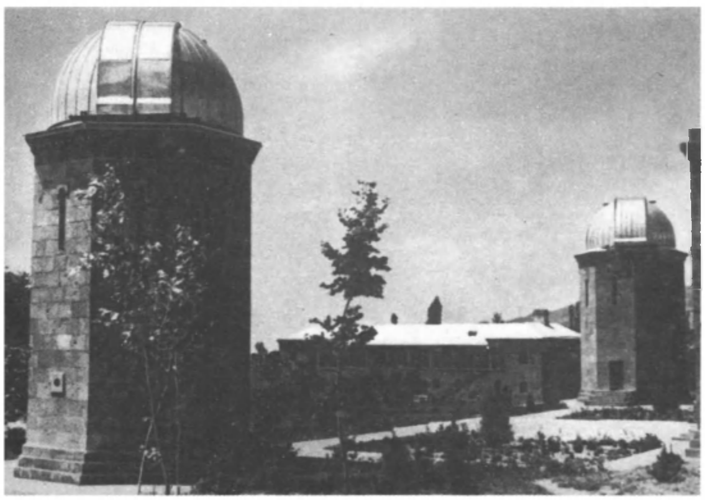
Using this and other spectral methods, and also absolutely different indirect methods, it is possible to determine the distances to stars lying beyond 300 light-years. The light of stars of certain distant stellar systems reaches us in a hundred million years. This does not mean at all, as some frequently think, that we perhaps observe stars that are no longer in existence. We should not say that "we see in the sky what does not exist in reality", because the overwhelming majority of stars change so slowly, that millions of years ago they were the same as they are now and even their apparent places in the sky change extremely slowly, although stars move through space rapidly.

Motion of Fixed Stars

The paradox is caused by the fact that in contrast to wandering celestial bodies – planets – the stars of constellations were once called fixed. Nothing can be fixed in the Universe, however. Two and a half centuries ago Halley discovered the movement of Sirius in the sky. In order to note the systematic change in the celestial coordinates of stars and their motion in the sky relative to one another it is necessary to compare precise determinations of their position in the sky, made at intervals of decades. They cannot be detected by the naked eye and during the history of mankind not a single constellation has changed its configuration appreciably.

With most stars no movements can be noted because they are too distant. A horseman galloping

The annual parallaxes of fixed stars: nearby star (left) and distant star (right).



The astrophysical observatory at Byurakan (Armenia, USSR).

on the horizon appears to be almost motionless, but a turtle creeping at our feet appears to move rather rapidly. The same applies to stars: we note more easily the motions of the stars closest to us. Sky photographs, which are easily compared with one another, are of great assistance. Observations of the position of stars in the skies were made long before the invention of photography, hundreds and even thousands of years ago. Unfortunately, they were too inexact to compare with modern observations for the purpose of detecting the motion of stars.

One faint star called Barnard's flying star for its noticeable motion amidst the stars in the skies moves, if you wish, not even with a turtle's pace, but far more slowly. It "flies" through the sky in an arc of 10″ in a year, i.e. moves by the apparent Moon's diameter (1/2°) in 200 years. If it is photographed with the large Moscow astrograph, which gives a large-scale photo, its annual change in position among the stars on the photograph will be less than one millimetre! However, in comparison with other stars this is actually a "flying" star. For most stars with already measured motion in the sky their displacement on photographs obtained using such large telescopes at intervals of tens of years is expressed in hundredths and even thousandths of a millimetre.

By determining the value and direction of the apparent angular movement of stars in the sky it is possible to compute how the arrangement of bright stars in a constellation will change in time. The motion of a great many stars has been determined, in particular, at the Pulkovo Observatory during the last century.

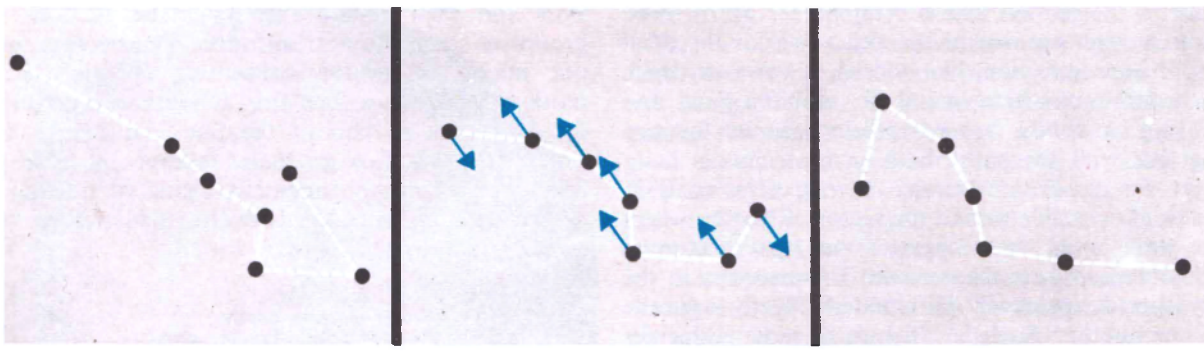
However negligible are the angular motions of stars in the sky, called *proper motions*, they correspond to an enormous velocity in space, if we recall the distance at which we see them. It must be remembered that the apparent angular motion of a star is dependent not only on the distance and velocity of its motion in space, but also on the direction in which it moves relative to us. A car travelling towards you on the road appears to move slower than a pedestrian who crosses your path at a closer distance. If a star moves at a right angle to the line of sight, its displacement in the sky corresponds to its velocity in space, which is easily computed knowing its distance.

Suppose that such a star moves 1″ a year and is located at a distance 2,062,650 times greater than the Earth from the Sun. Then, as we know, the length of an arc of 1″ (which is 206,265 times less than the radius of a circle) is a segment here 10 times greater than the Sun-Earth distance. This path, travelled by it in a year, can be used to compute its velocity in kilometres per second:

$$\frac{1{,}500{,}000{,}000}{365 \times 24 \times 60 \times 60} = 47.4 \text{ kilometres per second.}$$

We have still another possibility of studying the stellar motion–by measuring the displacement of lines in stellar spectra. The radial velocity, combined with measured proper motion and known distance (and only in this case!) fully determines the motion of a star in space, both in magnitude and direction. For example, if both velocities are equal, a star moves at an angle of 45° with our line of sight.

The combined efforts of many observatories, with Pulkovo and Simeiz occupying a leading position, enabled the proper motions of more than a hundred thousand stars and the radial velocities of seven thousand to be measured. Unfortunately, not all of these 7,000 are among the 100,000 for which proper motions have been determined. Stellar proper motions are determined by specialists on the



Proper motion: the migration of the stars of Ursa Major (from left to right) 50,000 years ago, today, in 50,000 years.

measurement of photographs who mostly hunt for stars with a greater apparent motion, whereas radial velocities are determined by spectroscopists, who prefer sufficiently bright stars, for which it is possible to photograph spectra with a large dispersion. The velocities of stars are typically tens of kilometres per second. The maximum velocity (583 kilometres per second) measured so far is that of a faint star in the constellation Columba.

"Traffic Control" for Stars

The first question to arise in the study of stellar motions is whether or not there is any regularity in the motions of stars. Are there paths in our stellar city along which the stellar population moves and is this motion controlled? If so, the role of a traffic officer is naturally played by the law of universal gravitation.

We want to know not only whether the stars move and what their velocity is, but what laws govern their motion. This required a vast body of observational evidence to be gleaned and our ability to analyze it.

As a result of such an analysis, it was found that certain groups of stars move in space in parallel paths and with identical velocity, being associated by mutual attraction and common origins. Examples are a group of faint stars around Aldebaran in the constellation Taurus called Hyades and five bright stars in Ursa Major.

In addition to such group motions, all the stars take part in a complex revolution around the centre of mass of our entire stellar system, but this will be discussed later in the book. Almost every star has its own proper velocity, like a gnat in a swarm of gnats carried along by the wind.

Where Are We Going?

Before studying the systematic motion of stars it must be clarified whether our own movement in space influences their apparent motion. Indeed, before clarifying the true arrangement of the solar system it was necessary to take into account the apparent motions of celestial bodies caused by the diurnal and annual motion of the Earth. It can be imagined that stars are fixed in space and our solar system moves among them. It would then seem to us that the stars were moving, giving way in places where we were moving and coming closer together in that part of the sky from which we were withdrawing, like the trees behind us during a walk in the woods. But in actuality stars are not like fixed trees. Each of them moves, and they are better to compare to small bow lights of boats in a large port that is entered serenely by a steamship during the night. They slide quietly in all directions, carrying their lights off into the darkness. Despite this chaos of motions we will note that the lights of the boats in front of us will seem to part as we approach and close behind as we recede. In order to detect the motion of the solar system it can be assumed that the motions of stars are disorderly and chaotic. We take sectors of the sky containing a sufficient number of stars and consider the mean motion of stars in them. Their chaotic motions in different directions are mutually excluded and there only

remains the motion that is common for all of them. Stars in such an area of the sky move for the most part in one direction, like a herd of cows in which each one grazes here and there, nibbling grass and moving randomly, but nevertheless moves forward together with the entire herd to fresh grass.

At our disposition we also have spectral analysis, making it possible to find the sector of the sky where the stars as a whole approach us with maximum velocity and where they recede. These sectors in the sky should apparently be situated directly opposite to one another. Such an analysis of radial velocities of stars can give us the velocity and direction of motion of the solar system, and of proper motions - its direction only.

These averaged systematic stellar motions, a reflection of motion of the entire solar system, suggest that it moves at a velocity of 20 kilometres per second in the direction of the constellations Lyra and Hercules. It is thus motion relative to comparatively nearby stars, taken as a unit, which manifests itself as a change in the apparent positions of stars, just as the apparent position of cows in the above grazing herd changes as you pass through it. In view of the relativity of motion, in the case at hand it is immaterial whether you make your way through the herd or whether they pass by you. In the same way we move relative to the stars.

The velocity of the solar system is of the same order as the proper velocities of stars. It is not to be feared that in flying toward the constellation Lyra we will collide and smash it into pieces. It would be more reasonable to fear that bullets, shot upward into a flight of planes, would destroy it. The constellation Lyra is only a direction in which a great many stars can be seen. The space between them is as ample as between the stars now surrounding the Sun. Light-years separate star from star. If you wish, try and compute in how many years we will cover half the distance to the bright star Vega in the constellation Lyra (ignoring its motion), assuming its distance to be 25 light-years and our velocity 20 kilometres per second.

The study of stellar motions advances by successive approximations, so to speak. We first assume that the motions are chaotic and determine the motion of the solar system. We then take into account its influence on the apparent motions of stars and then deduce the systematic motions of groups of stars. We next introduce a correction into our initial assumption concerning chaotic stellar motions and again, but this time more correctly, determine the motion of the Sun and iterate the procedure. It is thus gradually becomes possible to find our way in the apparent chaos of numerous stellar motions in our Universe and refine the picture drawn by the poet:

The celestial vault in starry glory
Cryptically watches us from its depths,
While we drift ever onwards, a flaming abyss
On every side encircling us.

(*Tyutchev*)

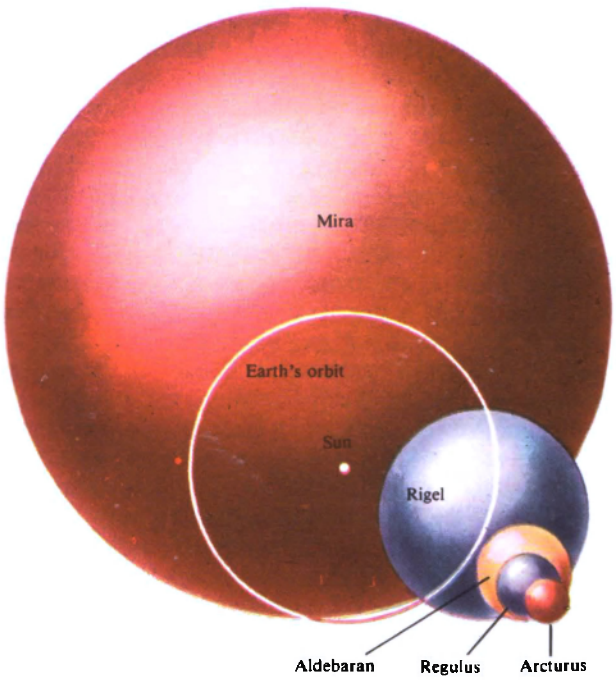
It is interesting to note that the seemingly arbitrary velocities of stars (like individual cows in a herd) are the greater the lighter are the stars themselves. The majority of heavy stars, like fat people, move slowly, and the light dwarfs are mobile, like children. However, there is a suspicion that in the stellar family it is the giants that are the children, not the dwarfs. But this is another story.

It is instructive that in a gas consisting of different molecules the heavier molecules also move more slowly.

Taking a Star's Measure

The dimensions of the planets are easily computed from their distances and the angular diameter of their visible disk. But how are we to measure a star when even through the largest telescope its disk cannot be seen, so small is its angular diameter? Even in a 5-metre telescope all the stars are visible as points. Here again physics helps us.

Since the stars radiate almost as an absolutely black body, the law of radiation of energy by them is known in different parts of the spectrum. If the temperature of a star and its luminosity are known, it is possible to compute the total energy emanating from the star. But for it, as a black body, theoretical physics is able to compute the total energy emitted by one square centimetre of its surface. According to the Stefan-Boltzmann law, it is proportional to the fourth power of temperature. If we divide the total energy emitted by the star, determined in this manner, by the energy emitted by one square centimetre of its surface, we obviously obtain the



Some stars shown to scale.

surface of the star. The star is a sphere, and knowing its surface a schoolboy can compute its diameter.

This method is entirely reliable, but as always in science, it is natural to want to find a way to check it. Such a method, applicable only to the brightest stars with a disk of the maximum angular diameter, was devised in 1920. It is based on the phenomenon called interference. In order to accomplish this Pease in the USA had to overcome a number of technical difficulties associated with the fact that even the world's largest telescope was insufficiently large for this purpose.

A solution was found by attaching a steel beam, 6 metres long, to the end of a 2.5-metre telescope (the largest at that time). A bogie with two large flat mirrors moved along the beam. The mirrors caught the light of a star and reflected it on the mirror of the telescope. Then the star image in the telescope constituted a tiny banded circle. At certain distances between the mirrors the bands on this circle disappeared and then the theory of interference made it possible to compute the angular diameter of the invisible disk of the star. Given the distance to it we can also compute its linear diameter.

The first star, whose diameter was measured in 1920 directly by an interferometer, was the bright red Betelgeuse star in the constellation Orion. In general the earlier measurements were a success for giant red stars, not necessarily nearby but such that their angular sizes as seen from the Earth were expected to be the largest. After about a dozen of such stars had been measured a long pause set in – the capacity of the instrument appeared to be inadequate. At last in 1956 in England scientists succeeded in measuring the diameter of Sirius, and in 1963 in Australia, the diameter of Vega. These were white stars, markedly smaller than the red giants, but the closest to us. In recent years also other techniques have been found to measure the angular diameters of stars.

We will cite the results of all these measurements and computations somewhat later. They show the extreme variety of the stellar dimensions. We note only that one of the largest among the known stars is the star VV in the constellation Cepheus. It is larger in diameter than the Sun at least 1600 times. There are stars far smaller than the Sun.

Star Pairs

The solitude of stars and their isolation from one another cannot be called a rule. Many of them form, probably for their entire life, pairs called double stars. They revolve around their common centre of mass under the influence of mutual attraction. It happens, of course, that two stars are sometimes visible close to each other in a telescope, but in actuality in space they have no relationship to each other. These are so-called optical binaries. In most cases, however, we deal with physical binaries, i. e. those which are attracted to each other.

This revolution around a common centre of mass was discovered for the first time by Herschel in England and confirmed by Struve in Russia.

By measuring the mutual position of double stars year in year out it is possible to determine the period of their revolution, which in most cases is extremely great and exceeds thousands of years. The shortest is about one year.

From such measurements we also determine the form of their orbits, but the true dimensions of the

orbits become known only when the distance is known. In actuality, the observations only give the angle at which it is possible to see the semimajor axis of the star orbit. Suppose that the semimajor axis of the companion orbit (the dimmer star) is visible to us at an angle of 2″ and the parallax of the star, i.e. the angle at which the semimajor axis of the Earth's orbit would be visible from the star, is 0″.2. It is clear that in such a case the semimajor axis of the orbit of the companion is 10 times greater than the distance from the Earth to the Sun.

Study of the mutual motion of double stars is unusually valuable for us for two particular reasons. First, it shows that the law of universal gravitation is also correct in the world of stars, far beyond the limits of the solar system. Second, it gives us our only possibility of determining the masses of stars. We recall that according to Kepler's third law

$$\frac{(m_1+m_2)P^2}{(M+m)T^2}=\frac{A^3}{a^3},$$

where m_1 and m_2 are the masses of our two stars, having the period P and a semimajor axis of the companion orbit A; M and m are the masses of the Sun and Earth, respectively; T is a year and a is the distance from the Earth to the Sun. If the mass of the Sun is assumed equal to unity, as also T and a, then by neglecting the tiny mass of the Earth,

$$m_1+m_2=\frac{A^3}{P^2}.$$

This simple formula then gives the total mass of the components of the double star, i.e. the members of this system. If we also determine from observations the distances x_1 and x_2 of stars from their common centre of mass, then as is known, $x_1:x_2=m_2:m_1$. From these two equations we then find the unknown mass of each of the stars separately.

It was found that unlike stellar luminosities and dimensions, the masses differ relatively little from star to star. The masses of giants are larger than the masses of dwarfs, but in general they all fall in the range from 40 times to 0.25 times the mass of the Sun. Only individual rare stars have masses attaining hundreds of solar masses. This uniformity in the masses of stars, together with the variety of their dimensions, suggests that the densities of stars should be extremely different. There is a clear dependence between mass and luminosity of stars (true, not all stars conform to this dependence) and it shows that only massive stars can have a great brightness, so that the mass of a star determines the relation between their temperature and dimensions.

But we will now return to double stars. Among them we find pairs that resemble twins, since the component stars resemble each other in all respects. There are star pairs which are just like a caricature, where an elephant and pug-dog are inseparable. In such cases the elephant is usually enormous, a bright but cool red star, and the pug-dog – its companion – is very small, faint, but hot and bluish.

Imagine that we are inhabitants of a planet that revolves around one of such stars. What amazing phenomena develop there in the sky. For example, from behind the horizon rises a red enormous circle, the sun, hundreds of times greater in apparent diameter than ours. Then a small bluish sun rises and gradually disappears behind the massive back of its patron, later to emerge again from behind it. Or else day comes, filled with red light, like sunset on Earth, but in place of night there sets in a blue day. The blue sun sometimes passes in front of the red sun and shines like a blue Bengal fire against a red background.

And what is it possible to see in a system of triple and even quadruple stars, where one of the stars or two are themselves systems of double stars, of different dimensions and colour! What extraordinary combinations of suns and what an interplay of colours should be there; how complex is the change between night and day, with a different number of suns in the sky, days lasting years, and sometimes even never passing into night-time.

Very close pairs do not reveal their nature, even in a telescope. The pair looks like a single star, but spectral analysis comes to our aid.

The penultimate star in the handle of the Big Dipper is Mizar, a second-magnitude star. Next to it, separated by 11′, there lies a naked-eye star noticed even by the ancient Arabs and called Alcor, which is the Arabic for "horseman". It is seen through a small telescope that Mizar itself consists of two similar stars separated by 14″ and Alcor appears to be very far from them. Mizar is a so-called visual binary star.

A great many pictures of the spectrum of the

brighter component of the Mizar pair – Mizar A – were taken at the Harvard Observatory in 1887-1889. Examining these photographs, Edward Pickering, director of the observatory, saw on some of them that the spectral lines were single, and on others double. When this was investigated in more detail, it was found that the spectral lines split periodically. More specifically, from time to time the lines split into two, the distance between them increases constantly, attains a maximum value, then decreases again; the lines again merge and then split again, undergoing all of these transformations with clock precision. The period of splitting is 10.25 days. But a displacement of lines corresponds to a change in velocity of the light source, argued Pickering, and the presence of two lines suggests the presence of two light sources, two stars. The spectra seem to be identical and superimposed on each other, since at Mizar A no two stars are seen in the telescope. If two superimposed spectra shift in opposite directions, their lines, first coinciding, will move apart. Sometimes several lines will split here, and two lines will sometimes merge.

The movement of spectral lines, first away from and then towards each other, must be caused by the motion of the stars, the possessors of these spectra. When one of them moves towards us, its spectral lines then move towards the violet end of the spectrum. At the same time, the other star recedes and its lines are displaced towards the red end. It can be seen that the velocities of stars simultaneously increase and decrease, and do so periodically. Pickering drew the conclusion, which appeared obvious: we have a close binary with an orbital plane close to our line of sight. Revolving around a common centre of mass and moving in their orbits, each of the stars periodically approaches us and then recedes. When both (always situated at opposite points of the relative orbit) move perpendicular to the line of sight, then, according to the Doppler principle, their spectra occupy a normal position and therefore merge together.

Soon thereafter a number of other stars were discovered with periodically doubling lines, also being *spectroscopic binaries*, as they are called.

This explanation of the doubling of spectral lines was confirmed in 1920 when by the use of an interferometer, employed for measuring stellar diameters, it was possible to measure the distance between the members of one spectroscopic binary. The negligibly small angular distance between them, measured with an interferometer but not perceptible directly in the telescope, coincided with that accurately computed from spectral lines. This star was Capella, and the angular distance between the component stars was at that time $0''.045$, a little less than the distance at which the two stars could be seen separately in the then largest telescope.

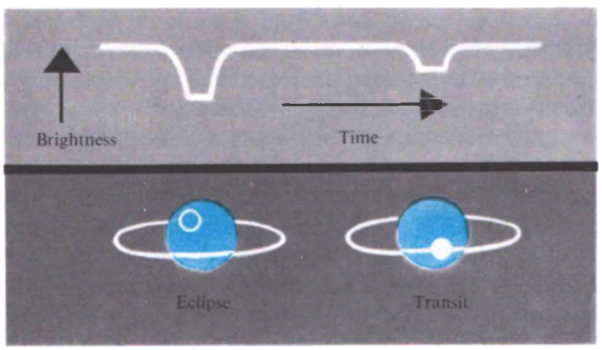
In most spectroscopic binaries the companion is too faint for the lines of its spectrum to be visible on the background of the bright spectrum of the main star. Then instead of the doubling of lines there is only a periodic oscillation of individual lines. It required a certain ingenuity to derive data on the orbit of the double system from observations of the radial velocity variations of a star. The periods of spectroscopic binaries are shorter – from 2 hours to 15 years. It may be that in a directly visible double star one member or even both are spectral binaries.

This also goes to prove that spectral analysis reveals all the secrets of stars and makes it possible to detect their invisible motions. The stellar spectrum is a passport that identifies it and does not permit it to hide even one of its secrets from our searching eye.

"Devil" Stars

The first devil star was discovered by the Arabs. This was β Persei, which they called simply "the devil" (El Gul). It surprised them in that, being usually a star of about the 2nd magnitude, it suddenly weakened to almost the 4th magnitude. It changed in Heaven, considered unchangeable. What could such a star be, if not a devil star, or the devil himself!

After a long train of historical events and the emergence of new centres of culture, several centuries later, a change in the brightness of β Per, El Gul, renamed Algol by the Europeans, was observed in 1669 in Europe. A century and a half later, the amateur astronomer Goodricke who was deaf-and-dumb from birth, discovered a periodicity in the brightness variation of Algol. Its period is 2 days 20 hours and 49 minutes. But during this period it retains a constant brightness for 2 days and 11 hours and then in the course of 5 hours it loses 2/3 of its brightness, so that 5 hours later it again returns to its former brightness.



The light curve of an eclipsing variable.

The strange and stubborn behaviour of the devil star was attributed to the fact that in reality there were two stars, one being far brighter than the other. They revolve around each other in orbit in such a way that from time to time the dimmer star partially conceals the brighter one from our view, producing periodic eclipses.

This explanation was confirmed at the end of the last century when it was found that Algol is a spectroscopic binary in which the spectrum of the faint companion is invisible, as might be expected. In this case at an eclipse the spectral lines occupy a normal position, i.e. the star at that time moves at a right angle to our line of sight (not towards or away from us), as should be the case. In addition, between the principal brightness minima there was found to be a small secondary attenuation of brightness corresponding to an eclipse of the faint star by the brighter one.

Many other double stars of this type, called *eclipsing binaries* or *algols*, have been discovered. Examination of the light curves together with spectral data, makes it possible to study these stars in a detail impossible in any other case. The "devil" stars are therefore the least mysterious of the stars now and nothing "devilish" remains in them for us except possibly the fact that they have been studied in "diabolic" detail.

As a result, we can find their form and dimensions in comparison with the Sun, the dimensions and form of the orbit and its position in space, the luminosities, temperature and masses of the stars and the character of the eclipses, and in addition in certain cases we can study the structure of their atmospheres almost in the same detail as for the Sun, although in the telescope these stars in no way differ in appearance from any other stars and appear as the same bright specks. Owing to their closeness, and hence the strong tidal forces, their configuration is elongated rather than spherical. They are elongated along the line connecting them and revolve "nose to nose", so to speak.

Portrait Gallery of Coloured Stars

We now will pass through a portrait gallery of stars and first glance at the typical faces of the ordinary inhabitants of the stellar Universe and then the portraits of certain stellar celebrities.

Here hang two portraits side by side, and in one of them we recognize familiar features, like those of our Sun: yellow colour, the spectral class G and a temperature of 6,000°C. Even the luminosity, mass, density and dimensions of this star are almost the same, but the inscription beneath the portrait tells us that this is not the Sun but the star α Cen A. In the depository of portraits of stars accumulated by astronomers there are whole stacks of exactly the same portraits, but with different inscriptions, indicating the name of the original, but if the inscriptions were taken away and a single name was attached to all the portraits nothing frightful would happen. These stars are so much alike that their own "children", the planets, could not tell the difference. The personality of our Sun is so commonplace that we could lose all interest in it if we did not need its daily heat and light.

Once again we arise our disappointed gaze to the portrait and we see that this is really only half the portrait of a star, and our closest neighbour in space, α Cen. Therefore, under this half the inscription reads α Cen A. Alongside, under the other half, is the inscription α Cen B. The fact is that α Cen is a double star.

From the second half of the portrait looks down a star of almost the same mass, a little lighter (by 15 per cent), but only 0.2 as bright. Whereas its companion on the basis of the colour of its envelope can be assigned, so to speak, to the yellow race of stars, this star must be considered a redskin. The deep orange colour of its surface corresponds fully to its spectral class K5 and a lower temperature of 4,000°.

The diameter of the star is 0.75 that of the Sun and the mean density is a little greater than that of water but less than that of the Sun. The general description of these portraits tells us that the period of revolution of these two stars is 78.8 years, a little longer than that of Uranus in the solar system, and the semimajor axis of the mutual orbit is 23.3 times larger than the Sun-Earth distance, i. e. again of the same order as the distance between the Sun and Uranus. However, here we have not the enormous Sun and a planet, but a pair of almost identical suns. The orbit of the companion, having an eccentricity of 0.51, is more elongated than the orbits of the major planets in our solar system and is more similar to those of short-period comets. The orbital plane is inclined to the line of sight by 11° only, so that the orbit is visible to us with great foreshortening. We know in what part of the orbit the companion is moving towards and in what part it is receding from us. Moreover, the entire system as a whole (its centre of mass) is approaching us at a velocity of 22 kilometres per second. However, it is moving at the same velocity (to be more exact, 23 kilometres per second) in a transverse direction, so that as a result the system α Cen is flying obliquely towards us at an angle of 45° at a velocity of 31 kilometres per second.

Under the portrait of α Cen we note a small portrait leaning against the wall. "You see", tells us the director of this "portrait gallery", "here we hang family portraits and I do not know whether we should hang together with the two α Cen portraits this portrait of Proxima Centauri, a star that is nearer than α Cen."

The fact is that it is only closer to us than α Cen by two light-weeks and is now moving in space in the same direction and with the same velocity as α Cen, but in the sky it is distant from it by more than 2°. This means that in space the distance between them is about one-thirtieth of the distance between the Earth and α Cen. You will recall that the latter distance is 4.3 light-years.

Undoubtedly, they feel some mutual attraction and their mutual motion is scarcely accidental. They possibly were once closer related and possibly they are connected by family ties, i.e. somewhere and sometime they were borne together, but it is difficult to be sure that this nearer star, Proxima, is a distant companion of α Cen and revolves around it in an enormous orbit with a period of thousands of years. It may well be that the common character of their motion is similar to the parallel linear motion of many stars in the Hyades and does not represent the motions of the components of a double star in a curvilinear orbit with a small curvature, such that cannot be detected due to the short period of our observations.

Look at the little portrait of Proxima Centauri, if you wish. It is small, because in our gallery the stars are painted to scale, if not "life size". This portrait is also one of the very great number of similar ones and is typical of stars called red dwarfs. True, among those known this dwarf exceeds in size only the star Wolf 359 and has been poorly studied, but nonetheless it is quite typical and it is not necessary to turn to other portraits just because the artist could not reproduce the features of the original with complete accuracy.

Proxima Centauri has a luminosity 1/15,000 that of the Sun and its mass is 1/7 that of the Sun. It is of a dark red colour, of spectral class M, with a temperature of only 3,000°. If we had it as our Sun it would illuminate us with a red light only 30 times stronger than the light of the full Moon. The diameter of Proxima Centauri is 1/6 the Sun's or only 1.5 times larger than Jupiter, the largest of our planets. The mean density of this red dwarf is almost 50 times greater than the density of water. If this matter were liquid, pieces of iron could float in it like corks. However, this is not a liquid, but a highly compressed gas, which in the star's interior has a very high density, but in the atmosphere is almost as thin as in the solar atmosphere. We now know that Proxima Centauri is a variable that flares rarely and disorderly, and that there are many such flare red dwarfs.

"Let's go to the courtyard", says the director, "and I'll show you a portrait of a star painted to scale. It is cool and "red-skinned" and of the same spectral class M. It is named Betelgeuse, or α Ori and its portrait will not fit in the gallery." In fact, after the portrait of Proxima Centauri, which you have looked at holding in your hands, you now must look way up in order to see the upper limb of the circle representing Betelgeuse. The portrait of our Sun was of the size of a human face, but the portrait of Betelgeuse rises 60 metres, reaching the 20th story!

The enormous body of Betelgeuse is 300 times larger in diameter and 27,000,000 times larger in volume than the Sun's. But it is only 15 times more massive than the Sun. Unlike the Sun, with its density 1.5 times larger than that of water, Betelgeuse is very light. It cannot be said that it is as light as down, because it is 1,500,000 times lighter than water, incomparably lighter than down and 1,500 times lighter than room air! If you filled the entire volume of the Bolshoi theater with such matter and then compressed this matter to the volume of a match box, you could freely put such a box into your pocket without fearing that the pocket would break under the weight.

Betelgeuse differs from Proxima Centauri, to which it is similar in colour, spectrum, and temperature by its luminosity (2,600 times higher than the Sun's), its giant size, its negligible density, and only to a slight degree by its mass. Betelgeuse is really a giant of the stellar world, to be more exact, a supergiant, and there are relatively few like it in the stellar Universe.

Turning to the colour, spectrum and temperature in search for a portrait of the Sun, we can rest our gaze on a portrait of Capella. It is also yellow and its other features are the same, but again it differs from the Sun in that it is a yellow giant, whereas the Sun is a yellow dwarf.

Like α Cen, Capella is, strictly speaking, not a star but a system of two stars. Spectrally, it is observed easily although it is at the limit of resolution for the world's largest telescopes. The period of revolution of the system is 104 days, 2 hours and 34 minutes; the orbit, almost circular, has a radius only a little smaller than the Earth's.

We would obtain a model of Capella, if in place of the Sun we put another yellow star, with a mass 4.2 times greater than the solar mass and a diameter 12 times larger than the Sun's; we then should replace the Earth by another sun with a diameter 7 times larger than the Sun's and a mass 3.3 times greater than the Sun's, and also 1,000° hotter. The new sun would be a yellow giant 110 times brighter than our Sun and its companion would be 69 times brighter. The mean density of both stars, yellow giants, would be twice as high as the density of the room air. The transverse velocity of the system would be 33 kilometres per second and the radial velocity 30 kilometres per second (recession), so that its total space velocity relative to the solar system would be 45 kilometres per second.

Yellow giants are also much fewer than yellow dwarfs.

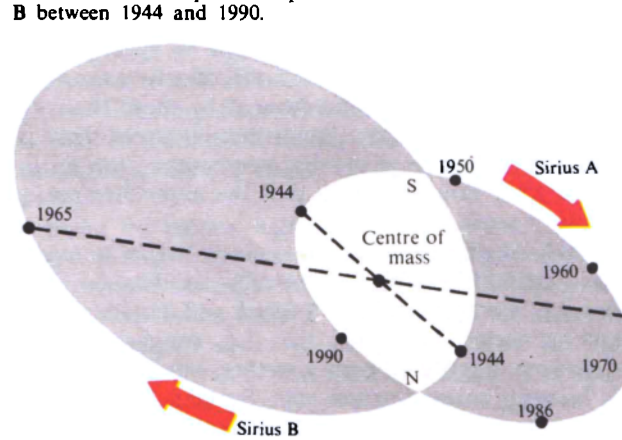
Portraits of White Stars and Their History

Among the typical portraits we still must look at the family portrait of Sirius. Each of the members of this star pair is a characteristic representative of white stars, encountered very frequently, but the Sirius pair surprises us with the disproportion of its members. Sirius A was given a laughably small companion by nature, but nevertheless it imitates its large neighbour in its manner of emitting a spectrum.

Really, their spectra are almost identical: A0 and A7, and the companion, being only two thousand degrees cooler than Sirius A, heated to 10,000°, is white as well. The mass of the main star is 2.5 solar masses, and of Sirius B is almost equal to the solar mass (to be more specific, 0.96), i.e. the masses of these two stars do not differ markedly. But here the similarity between the two ends.

Sirius A is a long-familiar type with a diameter almost twice as large as the Sun's and with a density 2.5 times smaller than that of water. Such are the characteristics of Sirius A, but the features of its companion, shining 10,000 times fainter, would seem completely improbable if the truthfulness of the portrait had not been repeatedly and authoritatively checked.

The observed and predicted positions of Sirius A and B between 1944 and 1990.



Anyway, it must be assumed on the basis of all the evidence available that it is only three times larger than the Earth (well and good) and its mean density is 30,000 times greater than the density of water. If this is so, we would need a lorry to carry a matchbox full of such matter, as its weight would equal that of 8 adults.

Who would believe this? Astronomers did not want to, but they were compelled, because all computations confirmed one another. Physicists declared on the basis of their successes in the study of the atom that such an enormous density is absolutely possible in the interior of stars. For this to be the case the temperature inside should be lower than in ordinary stars and the pressure of the above layers towards the centre should be enormous. In this case the atoms are destroyed and transformed into a mixture of nuclei and electrons, in no way connected to one another. The atomic nucleus is only 10^{-13} centimetres across, and the size of the atom can be determined as the diameter of the orbit of the outer electron shell (no less than 10^{-8} centimetres), i.e. the nuclei are a hundred thousand times smaller than the overall dimension of the atom. Consequently, the space occupied by the nuclei is far smaller than the space occupied by atoms, and with the adequate pressure, which exists in the interior of stars – white dwarfs, they can be squeezed far more closely together. Theoretically, crumbled atoms can be packed even more closely than in the companion of Sirius.

Ordinary atoms with an electron shell can be viewed as electric lamps with enormous paper shades, and crumbled atoms as these same lamps smashed into small fragments. Take these fragments and a box, which before would scarcely hold the lamp and shade, will now be able to hold a great number of lamp fragments, and so the box will become a thousand times heavier.

The companion of Sirius, initially considered a freak of the stellar world, a curio, was found to be a typical representative of a rather extensive family of stars, *white dwarfs*, which are difficult to discover only due to their general faint radiation. We can only note them when they are near us, like a beetle in a field, whereas bright stars, like cows in a field, are visible at a distance and in great number, although the cows are outnumbered by the beetles.

With respect to Sirius, it is worth mentioning that it has played an exceptional role in the development of human knowledge. The ancient Egyptians believed that Sirius controlled the floods of the Nile, which enriched the soil with silt. The flood was in the summer (after the falling of rain in the mountains of Ethiopia); its approach was "forecast" by a particular position of Sirius in the sky. Sirius, or the star of Isis, goddess of fertility, was also called Sopd, which the Greeks pronounced Sothis. When Sirius appeared before sunrise in the rays of the morning glow and the Nile flooded came the Egyptian New Year. They called to Sirius: "Divine Sothis calls the Nile to the beginning of the year. Great Sothis, shine in the sky and the Nile will come from its sources. Divine Sothis, produce the flooding of the Nile at its headwaters."

Sirius was the first star whose proper motion in the sky was observed by Halley, who compared his modern observations of the position of the star in the sky with ancient observations.

Almost a century and a half ago the example of Sirius demonstrated the strength of human genius, capable of creating not only a "visible astronomy", but an "invisible astronomy" as well. This was the remarkable history of the discovery of the companion of Sirius.

In 1834, studying the proper motion of Sirius, Bessel in Germany discovered that it does not move in a straight line (or rather, not in the arc of a great circle), but describes a sort of wavy line. It was not until ten years later that he concluded: the wavy path of Sirius is caused by the presence of some invisible companion with a period of revolution of half a century. The centre of mass of the system moves in a straight line, like any single star, but both bodies describe their orbits around it, so that the combination of orbital motion with translational motion gives the apparent path of Sirius its wavy appearance. Sirius and its invisible companion are always situated on different sides of their centre of mass.

The prediction was splendidly confirmed only on 31 January 1862. On that evening the American optician Alvan Clark tested a new 45-centimetre refractor he had constructed. This was the largest telescope then in existence, possessing excellent optical qualities. When he pointed it at Sirius, Clark saw a faint companion near it, precisely in the

position indicated by theory. Subsequent observations revealed that its period of revolution also coincided with the predictions. Thus, an invisible celestial body, whose existence was beyond question and whose daily position was known, became visible at last.

Sirius's companion is really not so faint, of the 7th stellar magnitude, but its nearness to the bright Dog-Star prevented it from being observed earlier and made it difficult to observe thereafter. Therefore, it was not until 1916 that it was possible to photograph its spectrum, which turned out to be similar to the spectrum of its bright neighbour. The assumption that it shines by reflected light led to the conclusion that it was of nonsensically large dimensions. Later a difference was found in the spectra of the two stars, which finally forced recognition that the companion shone by its own light. This was inevitably followed by the conclusion that it had an extraordinarily high density, which perplexed astronomers. In 1920-1924 the English physicist Rutherford found that atoms were complex systems consisting of nuclei and electrons and the Indian physicist Saha put forward the theory of ionization under high temperatures, and what earlier appeared completely unusual became perfectly natural.

By that time Einstein had developed the theory of relativity, initially met with distrust because of its shocking concepts. Einstein drew certain conclusions from his theory that could not be checked experimentally in the laboratory and required checking literally "on a universal scale" by astronomers.

A white "crumb" – Sirius's companion – was taken as an experimental rabbit or, if you wish, a white mouse. Its volume being small and its mass almost equal to the Sun's, it constituted what was needed for the experiment. As a result of the smallness of the size of Sirius B, the gravity force at its surface is a thousand times greater than on the Sun and almost 30,000 times greater than on Earth. A pendulum making one oscillation per second on Earth would make about 140 on the companion. A "day", i.e. 1,440 minutes according to the clock with such a pendulum, would be lived on a white dwarf in 10 earth minutes. According to the theory of relativity, under such conditions light oscillations (caused by oscillations in atoms) should occur appreciably slower than on Earth. Their wavelengths should be greater than in our laboratory, and the spectral lines of the companion of Sirius, coming from its surface, should be displaced towards the red end of the spectrum. The value of this shift should be the same as if Sirius's companion was withdrawing from us at a velocity of 20 kilometres per second.

But how to check that this shift actually occurs for such a reason? After all, the star may really withdraw from us at such a velocity and we assume the corresponding red shift to be due to Relativity.

Fortunately, the companion of Sirius makes it possible to distinguish such shifts. In fact, its velocity relative to us as a result of the orbital motion can be computed precisely for any time. Also, the motion of the centre of mass of the system of Sirius is well known from the observations of the spectrum of Sirius itself. It approaches the Sun at a velocity of 8 kilometres per second. If, both motions taken into account, there still remains some shift, the real motion is not the cause. To obtain a good photograph of the spectrum of the companion, outshone by a star 10,000 times brighter, was no easy matter. It was found that the spectrum of the companion really shows a line shift, exactly as required by the theory of relativity. The theory predicting the existence of the shift was confirmed thereby.

To what other scientific discoveries led the further study of the star of Isis and its strange companion? About dwarf stars and pygmy stars we will tell you in "Neighbours of the Sun".

Anatomy of Stellar Atmospheres

A bright distant single point in which nothing can be seen even in the strongest telescope – such is a star for an ordinary observer. For example, a star of the 3rd magnitude, ζ Aur, is one of those lamps lit in the evening by the angels which were held in respect by the superstitious people of the Middle Ages. For many years we did not see anything particularly special in it, but beginning in 1932 it became the most studied star. From 1932 through 1934, in two years, astronomers, like surgeons, fully dissected it, studying its entire anatomy, but instead of using scalpels they used the entire arsenal of instruments employed in the analysis of light.

In 1908 Campbell at the Lick Observatory

demonstrated that the spectrum of ζ Aur consists of two spectra, K3 and B8, superposed one upon the other, belonging to two stars invisible separately and making up ζ Aur. It was not until 1924 that the period of their reciprocal revolution – 972 days, almost 3 years – was established. Using all available spectrograms it was possible to compute the orbital elements and it was noted that one of the spectrograms differed greatly from all the others. The spectral lines in this spectrogram were unusually sharp. This was attributed to the fact that at the time when this spectrum was obtained the white star with the B spectrum was in eclipse – hidden from us behind the orange-red K star. It was thus concluded that eclipsing should occur periodically and ζ Aur therefore is an eclipsing binary, changing its brightness, a "devil star", an algol. The dates of future eclipses of the B star were computed (1926, 1929 and 1932) and plans were laid down to check these conclusions by observations.

These plans were not realized at once, and not until 1932 did two observatories at the appropriate time undertake measurements of its brightness and photograph the spectrum.

The prediction was brilliantly justified and a brightness minimum was discovered, lasting "40 days and 40 nights". Furthermore, an extensive atmosphere was discovered around the red K star, which due to the peculiarities of the eclipse was accessible to highly detailed study, and therefore in 1934 astronomers pounced upon this star. They hunted it out like reporters in pursuit of a visiting dignitary. Nevertheless, the most interesting period – when the eclipse was partial, i. e. when the B star partially looked out (although invisibly for us!) from behind the broad back of its stout and red-faced patron – was lost for sufficiently detailed observations. The partial eclipse lasted only 19 hours, of which many passed when the star was below the horizon and the others coincide with the day-time or with overcast weather. The total eclipse, however, lasts 40 days. How dissimilar is this to the solar eclipse, when the Moon, of the same apparent size as the Sun, within hours causes a partial eclipse and only for minutes, and sometimes only for seconds, blots out the Sun completely.

The red star ζ Aur is 293 times larger than the Sun and the white star is four times larger, so that the companion is 73 times smaller than the main star. When the companion hides behind the fat patron, it takes a long time to reappear.

The companion's orbit is like Jupiter's and the red star is almost equal in diameter to Mars's orbit and is 32 times more massive than the Sun, the companion however is 13 times more massive than the Sun. Its luminosity at a temperature of 15,000 °C is 400 times larger than the Sun's, and the luminosity of the red star at a temperature of 3,160 °C is 1,900 times larger. Together, they are 2,300 times brighter than the Sun. Like Betelgeuse, the red star has a very small density, corresponding to an air density at a pressure of 1 mm Hg (instead of 760 mm Hg of the standard atmospheric pressure).

The K star is surrounded by an extensive tenuous atmosphere through which the B star shines before hiding in the eclipse and also before emerging after the eclipse. Clearly, in the course of 2 or 3 weeks we can trace (from the spectrum) how the star B, passing behind the K, shines through the lower and denser reddish atmosphere. Its light is absorbed in this atmosphere, so that dark lines appear in the spectrum. The latter become stronger with increasing density and thickness of the atmospheric layers of the K star through which passes the radiation of the star B. By tracing the variation of the spectral lines it was possible to compute the quantity of atoms of different chemical elements in a column of the atmosphere at different heights above the surface of the K star. The atmosphere of the star was thus studied as if in profile, and almost better than the atmosphere of our Sun. It is to be stressed, however, that this star is visible to us only as a bright speck.

Changes in the brightness of a star as perceived by the eye are scarcely discernible – only 0.2 stellar magnitude, which is why earlier, without precise measurements, the variability of ζ Aur was not discovered. It was found that the farther we go into the violet region of the spectrum, the greater is the change of brightness there. In the UV light, invisible to the eye, with a wavelength of 3,780 Å, the change in brightness of the star attains 2.14 stellar magnitudes! This can be attributed to the fact that when there is no eclipse both stars shed light, but the hot white star alone is rich in ultraviolet rays and has appreciably fewer visible rays than the cooler red star, which is brighter to the eye. When the white star is in

eclipse, few visible rays but many ultraviolet rays are subtracted from the total light of the star, and so the change in their intensity is more conspicuous.

From observations of the spectrum before and after an eclipse it was possible to detect that the red star rotates with a period of 785 days in the direction in which it revolves.

As in the solar atmosphere, it is calcium that rises the highest above the surface of the star in its atmosphere. It attains a height of 233 million kilometres, i. e. it extends over the surface of the star by a distance 1.5 times longer than the distance from the Earth to the Sun!

Such is the "modest" information we receive about a bright speck lying at a distance that light takes 980 years to cover.

Stars Are Like Tops

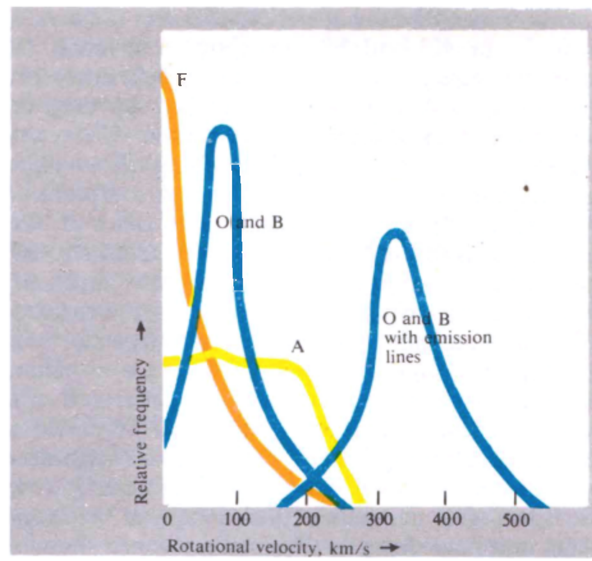
In a short story by the Russian writer Gorbunov, in olden days one of the attempts to popularize the teaching about the rotation of the Earth ended by a merchant's clerk punching his enlightener in the ear: "What? We live on a top, eh?"

But, alas, the globe keeps spinning, as it did millions of years ago, although ever slower and slower. Our neighbours, the planets, are rotating as well, with different velocities. His Majesty the Sun rotates too, if only sluggishly as compared with the planets.

But do stars rotate? The first to suggest the rotation of stars was Engels, who in his *Dialectics of Nature* maintained that all the bodies in space must rotate.

But how this rotation is to be found, if even through the most powerful telescopes a star appears as just a light point and its surface is not seen? But there is no "secret of nature" that, despite the idealists and agnostics, could not be revealed by human mind, sooner or later. As regards the stellar rotation, it has already been discovered. We owe this discovery to the American astronomer Struve and Soviet astronomer Shain.

We have already discussed the use of the Doppler effect to determine the velocity of a rotating body as applied to planets. Now by examining the line shape in stellar spectra we can also establish the rotation of a star and the velocity of this rotation. Not all the



The dependence of rotational speed on spectral type.

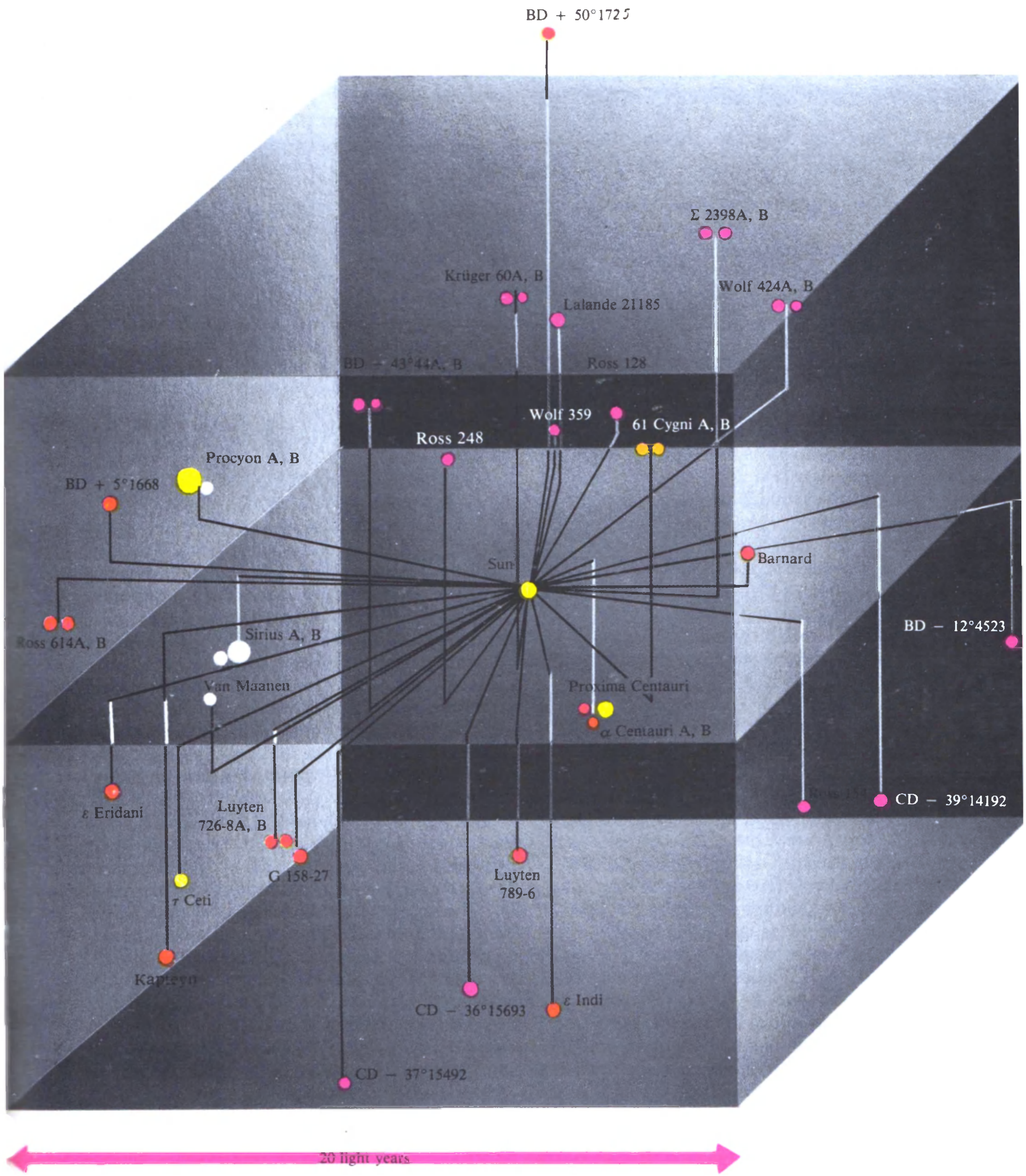
stars spin sufficiently fast for it to be discerned from their spectrum.

It turned out that the fastest tops are white stars. By way of comparison, the stellar equator rotates with a speed of 2 kilometres per second, and in these stars a point on the equator covers up to 500 kilometres in a second. Such stellar tops seem to be at the limit of their stability. Like a runaway fly-wheel, such a star may disintegrate. Really, as we shall see later, at least the atmospheres of these stars quickly scatter in space under the centrifugal force, which is possible even on the Sun. But the Sun rotates slowly, the force of gravity on it is larger than on white stars. Besides the temperature of the Sun's surface (6,000°) is far lower than that of white stars (10,000°), and hence atoms move about slower on the Sun. But the stars just described pale before pulsars, which take the cake among the stellar tops.

In the Neighbourhood of the Sun

If you are an inquisitive individual and settle in an unfamiliar city, your neighbours will be of considerable interest to you. To be sure, you are

The solar neighbourhood (within 15 light-years).


BD + 50°1725
Σ 2398A, B
Krüger 60A, B
Lalande 21185
Wolf 424A, B
BD + 43°44A, B
Ross 128
Wolf 359
61 Cygni A, B
Ross 248
Procyon A, B
BD + 5°1668
Sun
Barnard
Ross 614A, B
Sirius A, B
BD − 12°4523
Van Maanen
Proxima Centauri
α Centauri A, B
ε Eridani
Luyten
726-8A, B
CD − 39°14192
G 158-27
Luyten
789-6
τ Ceti
Kapteyn
CD − 36°15693
ε Indi
CD − 37°15492
20 light years

The Brightest 20 Stars and the Sun (according to B. Bok)

No.	Name	Apparent visual magnitude*	Spectrum	Absolute magnitude**	Luminosity	Distance (light-years)
1	Sirius	-1.6_d	A0	+1.3	23	8.7
2	Canopus	-0.9	F0	-4.6:	5,200:	180:
3	α Centauri	0.3_t	G0	+4.7	1.0	4.29
4	Vega	0.1	A0	+0.5	48	26.5
5	Capella	0.2_t	G0	-0.5	120	45
6	Arcturus	0.2	K0	0.0	76	36
7	Rigel	0.3	B8	$-6,2$:	23,000:	650:
8	Procyon	0.5_d	F5	+2.8	5.8	11.3
9	Achernar	0.6	B5	-2.6:	800:	140:
10	β Centauri	0.9	B1	-3.1:	1,300	200:
11	Altair	0.9	A5	+2.4	8.3	16.5
12	Betelgeuse	0.9_{var}	M2	-5.6:	13,000:	650:
13	α Crucis	1.4_d	B1	-2.7:	900:	220:
14	Aldebaran	1.1_d	K5	-0.5	120	68
15	Pollux	1.2	K0	+1.0	30	35
16	Spica	1.2	B2	-2.2:	600:	160:
17	Antares	1.2_d	M1	-2.4:	700:	170:
18	Fomalhaut	1.3	A3	+2.1	11	23
19	Deneb	1.3	A2	-4.8:	6,000:	540:
20	Regulus	1.3_t	B8	-0.7	140	84
21	The Sun	-26.72	G2	+4.8	1	—

* d-double; t-triple; var-variable.
** A colon indicates unreliability of the data cited.

familiar with all specimens of humanity – tall and short, fat and thin, blondes and brunettes, white collars and workmen, men and women, silent and talkative, gloomy and merry. Your neighbours will interest you in two respects. On the one hand, they will concern you as neighbours, as the "ambience" in which you live, and knowing them you will better define your own position in the world surrounding you. On the other hand, it is possible that certain statistical information about your neighbours will give you some ideas concerning what characteristics of the inhabitants of the city are typical, what type of people predominate in it and what is the rule and what is the exception. In approaching the city you possibly noted a small boy flying a kite, but it would be reckless to assume that the entire population of the city consists of small boys engaged in the same pastime.

It may happen that among your neighbours there will be no other more atypical representative of the population, such as the youngster fond of kites or a circus clown. But the existence of such residents can be taken into account and the presence of a clown (possibly at the other end of the city) can be learned of even from afar (from posters on billboards).

It is from these points of view that we are interested in the stars surrounding the Sun – the stellar Universe closest to us. Unfortunately, it is more difficult to become acquainted with the neighbours of the Sun than your next-door neighbour and especially so since their nearness in space is not so conspicuous.

To learn about the closeness of a star we must make careful and tedious measurements of photographs to determine its parallax and make sure it is large. In most cases we do not have unmistakable evidence of the closeness of stars and we cannot always select successfully those which after measurements actually will be close. Try and take a hundred stars at random and determine in succession the distance to each of them. After many years of work you will be convinced that not one of them has the right to be called our neighbour. Some of the stars will be at a measurable but great distance and for most of them the parallaxes will be less than the error that you cannot avoid in the measurements.

It was only natural that initially observers concentrated on the bright stars. They were easiest to study and it was postulated at first that those stars which appear to be the brightest are for the most part the nearest: their brightness seemed to be the result of their small distance. It turned out that this is not the case at all.

If we take 20 stars brighter than the 1st magnitude, we will find that they must be regarded as distant searchlights, not as nearby street lamps. Their enormous luminosity is what makes them the brightest in the sky and most conspicuous, although these searchlight stars are scattered far in space from us. The accompanying table gives data on these brightest 20 stars in our sky.

We can note the variety of distances and luminosities. The closest star, α Cen, is 4 light-years away and the most distant, Rigel, is 540 light-years away, whereas their apparent stellar magnitudes are identical. However, Rigel is approximately 23,000 times brighter than α Cen. But our neighbour has the same luminosity as the Sun and the least in the list of the 20 stars. This means that all 20 of the brightest stars in the sky are brighter than the Sun, mostly by tens and hundreds of times, and Rigel even by 23,000

times. The linear dimensions of all are far greater than the linear dimensions of the Sun. Particularly large are the red supergiants Antares and Betelgeuse, with diameters hundreds of times greater than the Sun.

More than half of the stars in the list belong to the hottest types of white stars of classes B and A, which are several times hotter than the Sun.

If we were to view this list as a typical example of the stellar population of the Universe, the Sun would not amount to much. Indeed, its luminosity is the smallest, it is small in size and even its temperature is below average. The Sun is definitely a very puny fellow. But would such a conclusion be correct?

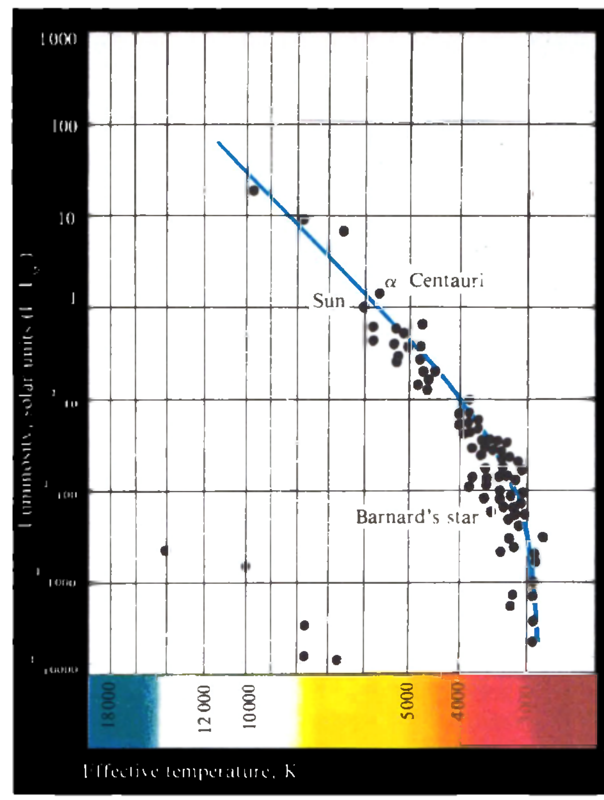
Neighbours of the Sun

We should change our opinion of the Sun if we study the nearest stars. By "nearest" we mean those that are situated within a sphere with a radius of 16 light-years, described around the Sun. Besides the Sun, we have now discovered about 50 stars in this sphere. These give us some idea of the density of the stellar population and predominant types of stars.

The following procedure has been used in recent years for detecting nearby stars from among the multitude. Luminosities of stars varying widely, their apparent brightnesses are unreliable indicators of their distance; whereas their apparent angular motion in the sky gives a truer indication. The velocities of stars in space are also extremely varied, but it is natural to expect that in general the more a star moves through the celestial sphere in a year, the closer it is to us, since apparent angular motion increases with decreasing distance. Experience has shown that by using this criterion we can actually detect many nearby stars.

Among them there are only four stars of about the 1st stellar magnitude: Sirius, Altair, Procyon and α Cen. Six others can barely be seen with the unaided eye, but only on a moonless night. The other stars in the list can only be seen in a telescope.

Thus, of the brightest twenty stars, 4 (20 per cent) are neighbours, but of 2,000,000 stars from the 9th to the 12th apparent magnitude there are only 20 neighbours, or 0.001 per cent! These stars of small brightness predominate among the nearby stars, and since there are extremely many such faint stars in the

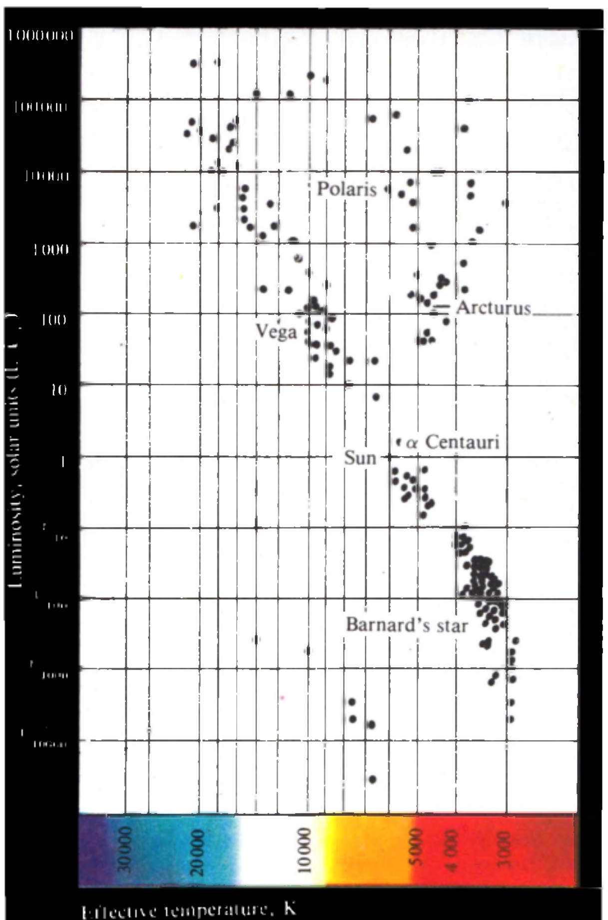


The Hertzsprung-Russell diagram for stars within 20 light-years.

sky, it is not surprising that it took much time to sort them out from this crowd. About half of them have been detected during the last 40 years.

The accompanying figures show the distribution of absolute magnitudes and spectral classes of the nearest stars. This is the *Hertzsprung-Russel* diagram, or the *spectrum-luminosity* diagram. All the stars are arranged on it approximately along the diagonal, which is an exceedingly curious fact to be discussed later.

Of the multitude of the nearby stars 12 are really multiple (10 double and 2 triple). Solitude among the stars is not as common as it was thought after the first binaries had been discovered. The latest discoveries have added to this list only stars of low luminosity. We are similar to fishermen who first catch the large ones and then the small ones.



'he Hertzsprung-Russell diagram for the brightest and nearest tars.

Iowever, the addition of new stars to our list with he passage of time convinces us that within the dopted limits of space we have already detected no ess than half of all the stars existing there. If there vere still many undiscovered stars in this region, they vould affect the velocities of those stars that we lready know.

Thus, in the stellar "living-space" considered not all ne inhabitants have yet been counted, but most (in ny case no less than half), and it is time to draw onclusions concerning what company our Sun keeps nd how typical is the sample. Among our neighbours nere are no very hot stars of class B and in general stars hotter than the Sun are few.

Perhaps still more characteristic here is the absence of giants, and especially supergiants, both in terms of luminosity and dimensions. The most frequent inhabitants of our defined region of space are the red dwarfs, cooler and smaller than the Sun, and with a far lower luminosity. They constitute half the stellar population.

The white dwarfs, like the companion of Sirius, are by no means exceptional freaks, as was once thought. Even in the small volume near the Sun we have discovered three, they are present in equal number with "normal" white stars such as Sirius and Procyon.

If we take into account the difficulty of discovering white dwarfs, it must be supposed that among the still undiscovered neighbours, in addition to red dwarfs, there are some white dwarfs as well. For example, up to 1935 only 3 white dwarfs were known (all near the Sun), but in the following years it was possible to determine the parallaxes of eight others, all more distant. The brightest in apparent magnitude is the companion of Sirius (the 7th magnitude), many others are of about the 12th magnitude.

The astronomer Luyten has now classified about 250 stars as white dwarfs. There are good reasons to assume that white dwarfs constitute about 1 per cent of the total number of stars in a unit volume. In 1963 Luyten discovered white star pygmies. The smallest of the known white and blue pygmies is, it seems, the hot star LP 768-500 Ceti. Its brightness is $18^m.2$. Its proper motion is enormous ($1''.18$ annually), and so the star must lie fairly close to us. If we take its distance to be 48 light-years, it will be 100,000 times fainter and 160 times smaller than white dwarfs of the type of Sirius's companion. Its diameter will be 100 times smaller than the Sun's, i.e. just like the Earth's. Its spectrum displays no lines whatsoever.

For the white pygmy LP 357-186 Luyten even assumed that it is half the size of the Moon and its density is around 200,000,000 grammes per cubic centimetre.

Of especial interest was the discovery of a pair of pygmies LP 101-15/16 $15^m.8$ with the annual motion $1''.62$. One member of the pair is a white pygmy, the other, red cool pygmy, such that half of its numerous spectral lines have not yet been identified. The spectrum will have to be studied further.

The speculations of Ambartsumyan, Zwicky and

others lead us to conclude that stars are possible consisting of neutrons or heavy elementary particles, hyperons. Being electrically neutral, these particles can be packed closer than nuclei and electrons in white dwarfs. As a result such stars can have a diameter of only several kilometres and absolutely unbelievable density–approximating that of atomic nuclei and even higher (about 10^{15} grammes per cubic centimetre).

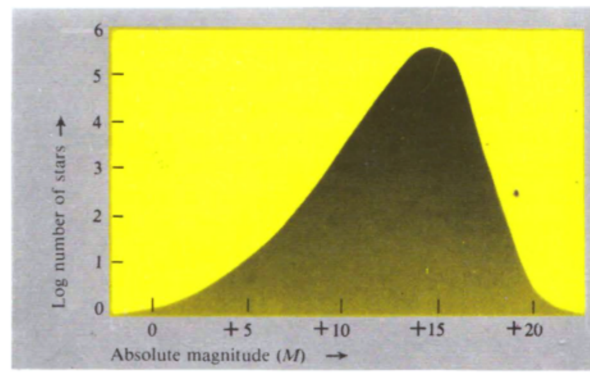
In theory, the neutron stars must produce intensive X-rays. Although this radiation has lately been found using high-altitude rockets, its source seems to be some other celestial bodies, rather than neutron stars. So far it has not yet been established if the above-mentioned superdense stars really exist as it is predicted.

Finally, the theory of relativity admits of the existence of extremely massive bodies of such density, achieved due to some catastrophic collapse, that no radiation escapes them. This fantastic hot body, stricly speaking, does not come under the heading of stars as it produces no radiation. It could only be found by its attraction exerted on other bodies. A more detailed discussion of them will be given later in the book.

Distribution of Stellar Luminosities

We have already seen from the Hertzsprung-Russell diagram that near the Sun, where the stellar population can be considered to be the most fully studied, it is the faint stars that form the majority. For a whole series of problems to be solved it is, however, important to know the precise form of luminosity distribution of stars. The luminosity grows with stellar mass, although much weaker.

To clarify these issues it is necessary to combine properly the results of studies of the bright stars with the results for the nearby stars. Otherwise, limiting ourselves only to the former data, we will not take into account the existence of dwarfs and within a sphere with a radius of 16 light-years we do not have a single giant. If we increased the radius of the sphere from 16 to 160 light-years, so that an adequate number of giants would be included (their number in a unit volume could then be estimated reliably), we would "lose" many dwarfs. The greater the volume of space we take, the greater the percentage of dwarfs in



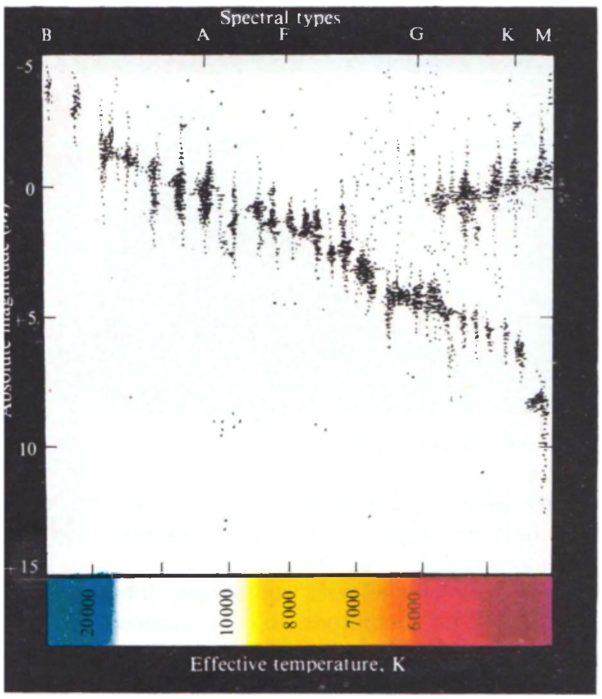
The luminosity function of stars.

it will still remain unknown to us, so that the truth can be approached only by combining the two methods of star counts.

Our data are reliable up to the 14th or 15th absolute stellar magnitude and the number of fainter stars can only be guessed, but in general it undoubtedly decreases, as shown by the curve representing the luminosity function. This curve gives the number of stars of the corresponding absolute stellar magnitude in a volume of 30 million cubic light-years. In other words, in our neighbourhood there is on average one star per 357 cubic light-years and the mean distance from star to star is about 9 1/2 light-years.

Luminosity function covers the stars irrespective of their spectral class and colour. Supergiants and even giants occur in the stellar population not with any greater frequency than professional clowns or persons taller than 2 metres are encountered among people. Most stars are dwarfs of the 14th-15th absolute stellar magnitude, whose luminosity is only 0.01 that of the Sun. The number of fainter stars undoubtedly decreases, and rather rapidly, although we do not know directly of stars fainter than the 18th absolute stellar magnitude.

If there were a very large number of dark stars, the observed motions of bright stars would be considerably different from what they are in actuality. Thus, these indirect considerations do not allow us to think that among bodies with masses of the order of the solar mass only a small part are sufficiently hot to shine like stars. Dark stars can exist, but they are undoubtedly fewer than the shining stars.



:he Hertzsprung-Russell diagram for many stars.

:ensus of the Star Population on the Iertzsprung-Russell Diagram

[he volume of space around the Sun that can be :onsidered quite well studied is too small and does ιot contain all the representatives of the stellar ›opulation. For example, it does not contain a single ;iant. It does not provide a full profile of the stellar ›opulation in general, like the population of your .partment house, not including a clown, does not orrespond fully to the entire variety of professions of he population of a large city. Therefore, striving to epresent all the types, but not endeavouring ιecessarily to know the number of all their epresentatives, we will consider the totality of stars ɔr which the luminosity is known (or the orresponding absolute stellar magnitudes) and their pectra (or corresponding temperatures and colours).

By constructing a diagram for several thousand tars we can see that they do not fill its entire area in disorderly fashion but are grouped within rather arrow bands.

The diagram reveals an extraordinarily interesting nd important fact, discovered for the first time by Hertzsprung of Denmark and Russell of the USA. Nature does not permit the existence of any stars of which our fantasy might conceive. For example, stars with a luminosity equal to that of our Sun but of a red colour (spectral classes K and M) do not exist. Stars have been discovered recently that are known as *subdwarfs*. Their luminosity is somewhat less than of dwarf stars of the same spectral class, lying on the main branch. Professor Parenago emphasized that they form a branch parallel to the main branch and believed that they are possibly even more numerous than the ordinary stars known so far. *Subgiants* have also been discovered. In terms of luminosity they fall between the dwarfs and the giants and are closer to the latter. The white stars, as we see, have only a determined luminosity, and this is extremely high. Moreover, yellow and red stars are only encountered either as dwarfs or as giants and the cooler (redder) the stars, the greater is the difference in luminosity between the dwarfs and giants. On the Hertzsprung-Russell diagram we see that in the neighbourhood of the Sun there is only the inclined branch of the diagram. It is called the *main sequence*, since it includes the overwhelming majority of the stars of our stellar system. A relatively small fraction of the stars falls on the branch of the giants, which is oriented horizontally on the diagram, and on the similarly oriented, but lying somewhat higher, sequence of supergiants. Our Sun is a star on the main branch, yellow, of spectral class G2 and with a luminosity normal for stars of this type.

The diagram shows first that only stars with a definite luminosity-temperature relation are encountered in nature. If stars have any other such relation they are apparently unstable, even if they exist, but we do not find them in the Universe. Second, the diagram gives the average absolute magnitude of a star in the main branch or a giant of a given spectral class. In short, if we know to what branch of the diagram a star belongs and what spectral class it is, we can use this diagram to determine the corresponding absolute stellar magnitude. It was found that for stellar systems of different structure and different age the Hertzsprung-Russell diagram has a form that differs widely.

The Hertzsprung-Russell diagram, that census of the physical behaviour of the stellar population, serves us as a constant reference.

3. Stellar Pulsations and Explosions

Cepheids – the Lighthouses of the Universe

Periodic variations of brightness are inherent not only in the algols, but also in other stars called *variables*. Of these, the hardest nut to crack were the *Cepheids*, which were so called after their typical representative, δ Cephei. Strictly periodically (a period of 5 days 10 hours 48 minutes), its brightness first grows by 0.75 of a stellar magnitude, and then fades slowly. It has also been found that as the star approaches its maximum brightness, its spectrum varies within the limits of an entire class, whereas the temperature varies within the limits of 800 °C.

Clearly, the brightness of the Cepheids varies not only for geometrical reasons, such as eclipse of one star by another, but for physical reasons. In fact, the physical behaviour of the star itself varies periodically, whereby varies the energy radiated by it, including the energy of light. The brightness variation is accompanied by periodic variation of the radial velocity of the Cepheids, a fact noted first by Byelopol'sky. As regards their brightness variation, it normally occurs within one and a half stellar magnitudes.

All of these variations can satisfactorily be explained if the Cepheids are treated as pulsating variables, which was first suggested by Umov and developed further into a theory by Shapley (USA), Sir Eddington (England) and especially by Zhevakin (USSR). Like rubber baloons, they alternately expand and shrink. It is this motion of their surface to and fro that is responsible for the radial velocity oscillations. As a star contracts, however, its temperature, according to the laws of physics, increases, whereby the spectral class becomes earlier, and the total brightness of the star increases anyway, despite its shrinking surface area.

The Cepheids seem to be unstable stars, which once perturbed by some push from inside oscillate like a pendulum. As the energy source that supports the oscillations depletes with time, the pulsations of the star must become weaker and finally die down. But nobody expected to witness this quite soon. So, fairly recently the first, and as yet only, gradual attenuation of brightness oscillation over four years was noted. The Cepheid RU Cam, first noted in 1899, has been changing its brightness by as much as a stellar magnitude with a period of about 22 days. By 1966 its variability almost completely ceased.

Different Cepheids have periods from 1.5 hours to 45 days, and after a period of shorter than one day they sharply jump to a period of more than two days.

All the Cepheids are giant stars of high luminosity, but they show a remarkable relationship: *the larger the period of brightness variation the larger its luminosity*. This brilliant discovery was made by Miss Leavitt of the Harvard College Observatory.

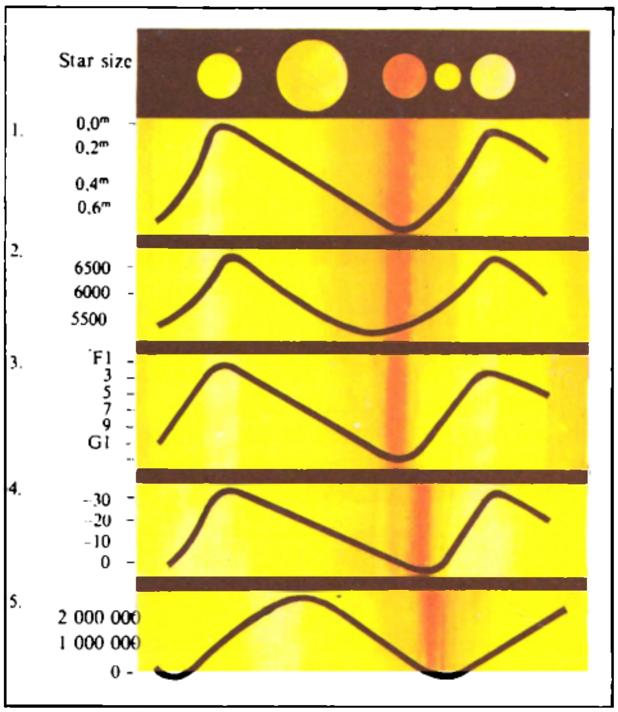
The Cepheids are, thus, sort of lighthouses of the Universe. With their high luminosity, they are seen at vast distances, which can be easily worked out given their period and apparent brightness. Accordingly, they are of help not only in our studies of the size, shape and structure of our stellar system, but also in studies of other galaxies.

Other Physical Variable and Flare Stars

Besides the Cepheids, coming under the heading of physical variables are many other stars. All of these vary not only in brightness but in spectrum as well, thus suggesting that the physical properties of these stars vary. However, with some of them, like with the Cepheids, brightness varies periodically, although not that regularly, and with others brightness varies semiregularly or even absolutely irregularly.

The most interesting stars are long-period variables. Their periods vary from 100 days to 700 days. The period from maximum to maximum is not always the same, also the light curve shape and maximum brightness vary somewhat. Almost in all of them brightness varies within the range of several stellar magnitudes, i.e. enormously; sometimes it changes several thousand times. They are sometimes called *Mirids* from the name Mira (Wonderful) given to the star o Ceti. At its brightest the star is well seen by the naked eye, being from the 4th to the 2nd magnitude. In about 330 days it reaches its minimum, the 9th magnitude, when it is only seen through the telescope. The maximum-to-maximum period may be either shorter (down to 320 days) or longer (up to 370 days).

Like the Cepheids, as the brightness of the Mirids vary, so do their temperature and spectrum, the latter showing at times the bright lines of hydrogen and other lines in addition to the dark lines. All the



The variation of the parameters of a Cepheid variable during its period:
1 - brightness; 2 - temperature, K; 3 - spectral type; 4 - radial velocity, km/s; 5 - increase in diameter, km.

Mirids are giant stars, they are cool, red, and tenuous. They also exhibit periodic variations of radial velocity, the maximum brightness corresponding to the maximum distance. It seems that, like the Cepheids, the Mirids vary in brightness due to pulsations, which although are less regular and complicated by both fluctuations of transparency of their atmospheres and periodic eruptions of hot gases from within a star to the surface, which accounts for the bright lines in the spectrum.

The other physical variable stars most commonly are also red giants and even supergiants with irregular, nonperiodic oscillation of brightness. Despite the irregularity, they can be subdivided into groups depending on the character of the irregularity.

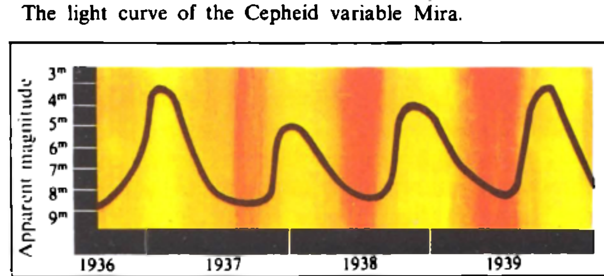
Some stars are characterized by constant irregular oscillations of brightness, others may have almost the same brightness for a long time, which only rarely weakens sharply but for a short period of time. Still others have irregular flares from time to time. There are stars that at times, sometimes for a long period of time, show a measure of periodicity. We cannot as yet account for this behaviour. Among this set of "epileptic" stars we will consider T Tauri and flare stars. They are attracting a great deal of attention now.

T Tauri stars are irregular variables, they are not red giants but stars of moderate or small luminosity, mostly of F or G spectral classes. Their brightness oscillations within one or two stellar magnitudes are accompanied by oscillations of brightness of wide lines observed in the spectra along with the absorption lines. The appearance of these bright lines is indicative of outflow of gases from the surface of such stars. In addition, their spectra at times show "continuous emission", a radiation with a continuous spectrum that partly masks the absorption lines and is obviously of nonthermal nature. It is caused by some unknown processes.

T Tauri stars occur as sparce, widely dispersed groups. Many of them have been found in tenuous regions of clouds of gas and dust, named diffuse nebulae. These stars are regarded as ones of the youngest, which may originate by the clustering of some regions of the nebulae. Some of the T Tauri stars are surrounded by their "own" tiny gaseous or gas-and-dust nebulae, which goes to support the above assumption.

Apparently, similar to T Tauri stars in nature are the so-called Herbig-Haro objects. The two astronomers have discovered in the region of dark dust nebulae some extremely weak stars surrounded by tiny nebulae. Their spectra with bright lines are similar to those of T Tauri stars and have, besides, the lines of exceedingly tenuous gases. Their brightness

The light curve of the Cepheid variable Mira.



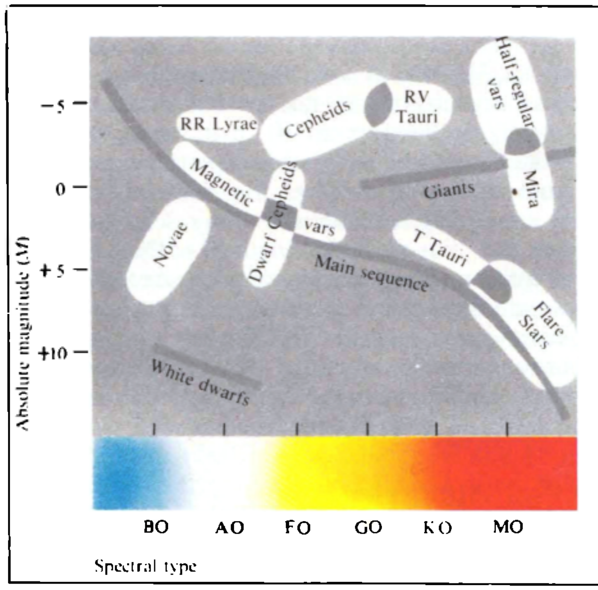
varies in an irregular manner, but their luminosity is far lower than that of T Tauri stars. Some of such objects were discovered by Herbig where earlier pictures showed nothing, as if they emerged during a period of as short as several years. But scientists are as yet not completely sure of this, since the earliest picture of these places has only been taken relatively short of the discovery. It may well be that earlier still the objects were seen too, and during the period the first picture was taken they only showed a temporary deterioration in their brightness. It is possible that Herbig-Haro objects are an initial stage of the emergence of T Tauri stars from diffuse matter by compression. This does not contradict the fact that some gas can flow out of their surface, the causes of which phenomenon are as yet unclear. In any event, the Herbig-Haro objects mostly resemble an incipient star.

Flare UV Ceti stars are red M dwarfs of very low luminosity. Their spectra have emission lines of hydrogen, helium, calcium, and iron. Most of the time these stars are almost constant in their brightness, but at times a star would quite unexpectedly flare, its brightness would intensify several times during a few minutes, but already in several minutes it returns to normal. It appeared that during their flares the UV Ceti stars beam out powerful radio emission into space.

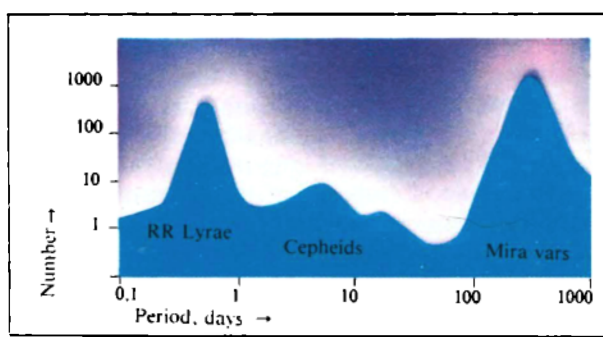
By the nature of amplification of ultraviolet radiation and other changes in their spectrum, the flares of UV Ceti stars are similar to chromospheric flares on the Sun. These latter are also accompanied by flashes of radio-frequency radiation. But in this case the flares of the stars are accompanied by a 100-1,000 times larger energy release than is the case with chromospheric flares. Evidently, for some reason or other, the UV Ceti stars release at times clouds of hot gases that contain many relativistic electrons, which, when decelerated in the magnetic field of the star, produce the radio emission.

Owing to their low luminosity we only see those of the stars of this type that are the closest to us, and owing to the flares being so short they are noticed but rarely. Therefore, by 1967 only 11 such stars have been found, but in a decade already about 500 have been recorded.

A good example of international cooperation in astronomy is the "flare star service". In some

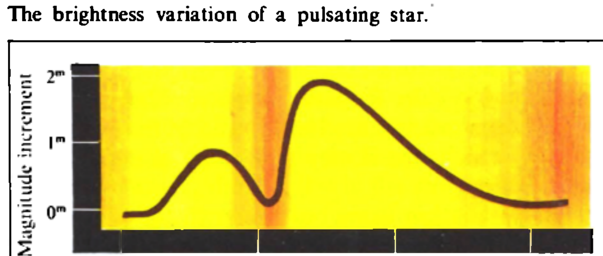


The positions of variables of different types in the Hertzsprung-Russell diagram.



The distribution of observed periods of variables.

The brightness variation of a pulsating star.



observatories, say in the USSR, they watch a flare star for many hours (maybe it would flare), whereas in Britain or Australia it is watched by a radio telescope. Such an organization of studies enables the cases of successful observations to be made more frequent.

In the Soviet Union, besides experts, a great army of amateurs are involved in the studies of variable stars, among them there are many schoolchildren. Precise and frequent determinations of the brightness of variables may yield results of great scientific value.

The achievements of the Soviet science in research of variable stars account for the fact that Moscow astronomers were entrusted to coordinate international studies in this field: the naming and cataloguing of variables, and so on.

Swollen Atmospheres

As you know, the spectrum is the passport of a star. It tells us about its physical state, provided we know how to decipher it. It has only been during the last two decades that we have been able to interpret the spectra of the stars we will now discuss, and in certain stars the spectrum is very unusual. In the overwhelming majority of stars we have discussed up to this point the spectra have been the same as that of our Sun – continuous spectra with dark lines. Bright lines only appear from time to time in the spectra of long-period variables and apparently are evidence of periodic powerful eruptions of hot gases onto stellar cooler surfaces.

But certain hot stars of the spectral classes A, B and O have spectra with individual narrow bright lines and in most cases certain dark lines are bounded from the red end of the spectrum by adjacent bright lines. It is remarkable that this is encountered in hot stars, and in general the hotter the stars, the brighter are these lines in their spectra and the greater the percentage of stars that reveal them. The scarcely noticeable bright lines adjacent to certain dark ones can be found in such exceptionally bright A0 stars as Deneb (α Cyg). In B stars, which are hotter, these lines are seen much more clearly, and in the spectra of O stars, which are hotter still, they are the most conspicuous.

The explanation came only recently. It was found that such stars have exceptionally extensive, swollen envelopes. These so-called extended atmospheres are characteristic of the hottest stars, having the greatest luminosity.

The reversing layer, as is well known, absorbs the light coming from the hotter, underlying layer of the star or photosphere, but *itself* emits the *same* wavelengths as it absorbs. Absorbing the light which reaches it from below in the star, it then re-emits in all directions and therefore we see only part of the light having the wavelength absorbed by the reversing layer. However, in the adjacent wavelengths to which the reversing layer is transparent, the light reaching us is not attenuated, and, as a result, in the visible spectrum of a star we can see a dark line of a particular wavelength. At the solar limb, where there is no photosphere behind the reversing layer giving a continuous spectrum, we perceive the radiation of the reversing layer itself and observe a spectrum of bright lines.

On the Sun the reversing layer and chromosphere are very thin in comparison with the solar globe itself, like the shell on an egg, and those parts of them that are projected beyond the solar limb are very narrow, yielding a weak emission spectrum. This spectrum can only be seen without difficulty at the time of total eclipses of the Sun when it is not "submerged" by the continuous spectrum of the photosphere. The latter is obtained from skylight, i.e. from the Earth's atmosphere, scattering the light of this photosphere. Skylight enters the spectroscope slit together with the light of the chromosphere, because the chromosphere is visible against the sky. At times when there is no eclipse the bright spectral lines of the chromosphere are too faint against the background of the bright photospheric spectrum and can only be seen with the greatest difficulty.

But in the case of stars whose chromosphere has a thickness comparable to the radius of the star, the radiation of the thick chromospheric ring is comparable to the radiation of the chromosphere and the bright lines become visible against the background of the continuous spectrum. In stars with such "swollen" atmospheres their extended chromosphere can be compared to the shell of a green hazel-nut.

Hot stars often have swollen atmospheres because the hotter they are, the greater is the percentage of ultraviolet rays in their light and it is ultraviolet rays that cause the strongest light pressure on the atoms in their atmospheres to counteract gravitational pull,

which attracts the atmosphere to the surface of the star, decreases the weight of the atoms and therefore probably makes it possible for them to rise to a greater height above the surface. The atmosphere therefore expands, becomes extended and swollen.

However, very swollen atmospheres are also encountered in very cool stars.

Some astronomers think that the rapid rotation of a star greatly facilitates the formation of an extended atmosphere. In fact, hot stars rotate on their axes more rapidly than others. Then the centrifugal force weakens the attraction of atoms by a star, and they can escape more easily from its surface. If this is actually the case, swollen atmospheres will be extended to the greatest degree in the equatorial plane of such a star, where the centrifugal force is greater. The star itself can be almost circular, but its atmosphere can be turnip-shaped. In certain cases a sort of Saturnian ring can form around the star, although gaseous. Of course, this cannot be seen in a telescope, because of too great a distance, but such an assumption would explain a great many peculiarities in the spectra of certain stars with bright lines.

The dimensions of the extended chromosphere and the brightness of spectral lines are dependent not only on the temperature of a star and the rotational velocity, but also on the luminosity of the star and gravity at its surface. The latter is dependent on the relation between the size and mass of the star.

Stars with Escaping Gas

The stars described above hold extended and swollen atmospheres, the same as the Earth holds its atmosphere. But in the collection of stellar spectra it is possible to trace a transition from spectra with individual thin *lines* to spectra containing individual unusually broad bright *bands*, side by side with dark lines and even without them. If there are dark lines, each of them is usually adjacent to a bright band in the direction of the violet end of the spectrum and is very greatly displaced from its normal position. The middle of a bright band, however, occupies a nearly normal position, corresponding to the line of a particular chemical element. The emissions of neutral atoms of hydrogen and helium and the ionized atoms of nitrogen, carbon and oxygen most frequently have the form of broad bright spectral bands.

If the dark lines are taken to be displaced from their positions as a result of the Doppler effect, i.e. due to the motion of the corresponding gas in the stellar atmosphere, it turns out that these dark lines at the border of bright bands have been formed by gases moving towards us at a velocity in some cases attaining 2,000 kilometres per second!

Stars that could be assigned to O stars on the basis of their spectral lines, but have broad bright bands in the spectrum, are called *Wolf-Rayet stars* after the two French scientists who first discovered and described them in the 19th century. The nature of these stars has been interpreted only recently.

Stars of this class are the hottest of any known. Their temperatures, measured by the author, fall in the range from 40,000 to 100,000 K. The hottest of ordinary O stars (without bright spectral bands) have a surface temperature of 30,000 K only.

Such enormous temperatures are accompanied by such powerful radiation of a flux of ultraviolet rays that light atoms of hydrogen and helium, and at very high temperatures also atoms of other elements, apparently unable to withstand light pressure from below, escape upwards at an enormous velocity. The velocity of their motion under the light pressure is so great that the attraction of the star is unable to hold them. They break loose from the surface of the star and almost without restraint whirl through universal space, forming a sort of atomic shower, but directed upwards, rather than downwards. Such a shower would scorch everything living on planets which might surround these stars. The continuous shower of atoms escaping from the surface of the star forms around it an atmosphere, which is continuous but constantly dissipating into space.

The gas atoms making it up are replenished continuously by the star. In the spectrum of such a star in place of a bright line we see a broad band, made up of a great many lines that have been displaced from normal position and have merged with one another. Each of these has been formed by atoms flying at some angle to the line of sight. The greater the angle, the smaller is the projection of the velocity of the atom on the line of sight and only its value (and not space velocity) causes the shift of spectral lines in accordance with the Doppler

principle. Because of the enormous size of the radiating atmosphere of a Wolf-Rayet star, hidden beyond the body of the star are only a few atoms, receding at maximum velocity. However, the atoms approaching us with maximum velocity are projected on the star and it is assumed therefore that they give rise to the dark absorption line. It is obviously displaced towards the violet end of the spectrum.

How long can a Wolf-Rayet star emit gas? It has been estimated that in one year a Wolf-Rayet star ejects a mass of gas equal to approximately one ten-thousandth or one hundred-thousandth of the solar mass. And will it eventually become exhausted? Will such a star emit gas until there is none left? The mass of a Wolf-Rayet star is on average a dozen of solar masses. With the gas escaping at such a rate a Wolf-Rayet star cannot have a lifetime longer than 10^4-10^5 years. Some independent data make it possible to assume that in actuality stars in such a state exist no longer than ten thousand years, and probably even a shorter time. With a decrease in their mass to a certain value their temperature probably decreases, the escape of atoms stops as does the self-destruction of the star.

So far we know of the existence of only about a hundred such self-destroying stars in the entire sky. It appears that probably only a few of the most massive stars attain such high temperatures in the course of their development that a loss of gas begins.

Most of Wolf-Rayet stars (the average mass is about 10 times larger than the Sun's) are very close spectral binaries. Their companion is always a massive and hot O or B star. Many of these stars are observed as "diabolic", i.e. eclipsing binary stars, which periodically blot each other from us.

Stars That Cast Off Their Envelopes

Stars of the Wolf-Rayet type became known to us recently, but they are closely related to stars of another type, known for 2,000 years, which have remained the most mysterious of all the stars ever seen by mankind.

Imagine the world as seen by the ancients – a crystal-clear celestial sphere with its mysterious star-lights fixed since the day of creation. Fixed and unmodified points of fire, shining on us as on our distant ancestors, indestructible groupings of stars of the constellations Scorpius, Auriga, Ursa Major and others.

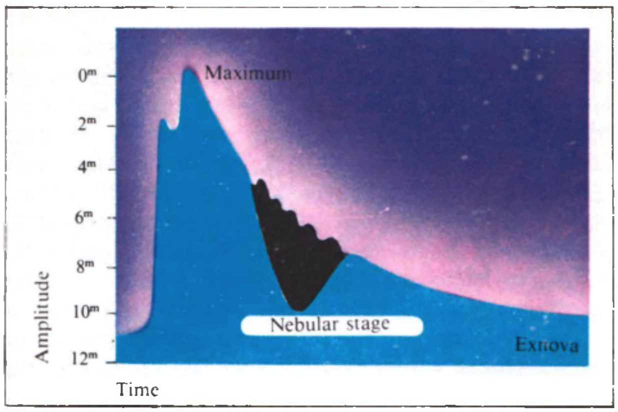
And suddenly in the 2nd century B.C. the great scientist of antiquity Hipparchus noted in Scorpius a bright star such as neither he nor his ancestors had seen there. What was it? A new act of creation, taking place before our eyes, or a correction to the already created, unchanged world? The new star in Scorpius did not shine very long and after losing brightness was lost from view. The stunned Hipparchus decided to undertake a listing of stars in the entire sky and record precisely their positions and brightnesses so that generations to come could determine whether new stars would again appear in the sky or whether those long known would disappear. Thus it was Hipparchus who compiled the first known star catalogue. On the basis of his example others were created; their scope was considerably broader and constituted the foundation of astronomy.

Cases similar to that observed by Hipparchus were also noted later. They were noted by Chinese and other chronicles. The first such cases to be described in Europe were in 1572 and 1604, on the eve of invention of the telescope. And in all cases the end was the same – the suddenly flared star, this new celestial body, had only had a short lifetime and disappeared from sight after a few months, becoming increasingly fainter each day. The common opinion was that the new stars, as late creations in an already finished and perfect Universe, were unstable and therefore disintegrated rapidly.

Modern methods of stellar research have made it possible to unravel to a considerable degree the secret of nova explosions. The last observed novae, shining for some time as stars of the first magnitude, were in 1918 in the constellation Aquila and in 1934 in Hercules. Many fainter novae have also been observed and it has been computed that in our stellar system – the Galaxy – no less than 200 novae explode each year, although we only observe the brightest of these and by no means every year.

Many novae have been discovered by nonspecialists and amateur astronomers. For example, the nova in Pictor was discovered in 1925 by a postman, the nova in Perseus was discovered in 1901 by a schoolboy in Kiev and the nova in Corona Borealis was observed first by a railway worker, and so on.

Collections of sky photographs stored at



Schematic light curve of a nova.

observatories (so-called glass libraries because they consist of glass-plate negatives) have helped to establish the following fact: novae are not at all new... They existed earlier, but as inconspicuous faint stars. When we note a bright nova, it turns out to be a faint star that suddenly has grown in brightness. The outburst occurs very fast, usually within two days. During this time the star becomes 11 stellar magnitudes, and sometimes even 14 stellar magnitudes, brighter. This corresponds to a brightness increase by a factor of 25-400 thousand times. If we conceive of the brightness of a nova as represented by a column of an appropriate height and assume that before the outburst the column was 1 centrimetre high, the brightness at the maximum would be represented by a column up to 4 kilometres high. It would be higher than most clouds and almost level with the highest peaks of the Alps and Caucasus.

In other words, a nova explosion is similar to a candle burning on a table that is suddenly transformed into a searchlight. Of course, in this analogy we must remember the scale of the phenomenon. The star, however faint it may have been before the explosion, was nevertheless a star, not a candle. Moreover, it has been established that the luminosity of novae before an explosion in most cases is of the same order of magnitude as the luminosity of the Sun. Just imagine what would happen if our Sun decided to flare up! If its radiation increased tens of thousands of times we would be not only blinded, but also burned up.

Immediately after a nova has attained its maximum it begins to fade, first rapidly and then more slowly; after several years the star returns to the same brightness that it had before the outburst. During the abatement of brightness there are frequent secondary explosions, but despite these common features it is impossible to find two such stars for which the light curves are completely identical.

Such an enormously large jump in brightness in itself is indicative of a catastrophic origin, but the spectral data give a still more interesting picture whose details have only been interpreted successfully in the last two decades.

Shortly before its maximum brightness a nova has usually a normal stellar spectrum with narrow bright lines on the border of dark lines, the appearance of which is characteristic of supergiants, i.e. large stars of enormous luminosity and extended atmospheres. This is also confirmed by data on its luminosity derived from estimates of distance to the nova. At the maximum of its luminosity its intensity exceeds that of the Sun by tens and even hundreds of thousands of times. For a short time the nova exceeds luminosity of all other known stars. The dark lines at that time are shifted towards the violet end of the spectrum, the shift corresponding to the velocity of approach towards us – several hundred kilometres per second. However, the temperature of the star at this time is almost constant and not too high: 8,000-10,000 K. The spectral class is A or F.

All this suggests that a sudden increase in brightness is caused by a sudden increase in the dimensions of star envelopes. Its outer layers, together with the photosphere, reversing layer and chromosphere are blown away like a soap bubble. They are carried off in all directions from the centre with a velocity of hundreds of kilometres per second, but we see only those parts that face us, i.e. only those that are approaching us (the remainder is hidden by the body of the star). For this reason the spectral lines are displaced towards the violet end.

A surprising phenomenon occurs immediately after the brightness maximum. The dark lines of the spectrum of the nova are replaced by broad bright bands, at the violet edge of which there is a well-defined dark line displaced from its normal position by an amount corresponding to the velocity

of approach – a thousand kilometres per second or more.

The explanation for this phenomenon here is the same as for Wolf-Rayet stars: expansion in all directions of an extended atmosphere, transparent to its own radiation, so that we perceive light from the receding parts, forming the so-called red half of the bright spectral bands. But in this case the principal role is played not by the continuous loss of atoms from the surface of the star, but by expansion of the atmosphere, which is torn away from the star at the time of the maximum. The reason for this detachment is a sudden increase in the velocity of atoms of the atmosphere under the influence of increasing light pressure.

Thus, at its maximal brightness the star inflates like a soap bubble and sheds its veils. These veils expand in receding becoming more tenuous and transparent, the naked star showing through them.

Imagine once more that our Sun suddenly began to swell like a bubble. We would be burned up like pieces of straw because the swelling of star causes the star at the maximum to have a diameter greater than the diameter of the Earth's orbit. We would be within the star even before it shed its envelope. But as will be shown, *such a catastrophe cannot occur to our Sun.*

Beginning with the time of the maximum, there are continuous and strong changes in the star spectrum, these are of the greatest interest for the specialist. But you are no specialists and I will not tire you with those details with which I have been concerned for many years. Suffice it to mention that the study of spectral changes leads to the conclusion that in the course of time a nova is heated more and more and eventually reaches a temperature of about 60,000-70,000 K and acquires a spectrum of the Wolf-Rayet type. If we were not familiar with the entire history of this star we might think that it was an ordinary Wolf-Rayet star. However, ordinary stars of the Wolf-Rayet type are almost a thousand times brighter and this forces us to use caution and not include them in a list of "former" novae.

But what happens to the gaseous envelope cast off by a star at the time of its brightness maximum? Upon expanding it is carried in all directions into space, constantly receding from the star. If you doubt the correctness of such an explanation of phenomena in the spectrum observe a nova several years after an explosion using a large telescope. By that time the envelope, which has become less bright but keeps expanding, has become sufficiently large for it to be seen directly in a telescope, even at the enormous distance from the Earth. We can actually see such nebular envelopes around the former novae of 1901 (Perseus), 1918 (Aquila), 1925 (Pictor), and 1934 (Hercules). From year to year we measure a continuous increase in their dimensions, occurring with the velocity determined earlier from a study of the spectrum. Such a nebula has initially the form of a tiny spot; the spot then increases and is transformed into a ring, at whose centre we can see a faint star, the former nova. Computations reveal that the bulk of the envelopes cast off by the star has a mass tens or hundreds of thousands of times less than the Sun's. The cast-off envelope consists of hydrogen, helium, nitrogen, carbon, oxygen and other gases. In form, chemical composition, and physical state the gaseous nebulae formed by novae resemble the small nebulae encountered in various parts of the sky, formerly inaptly called planetary nebulae due to their external similarity to the greenish, slightly luminescent disks of the planets Uranus and Neptune.

The similarity between nebulae caused by novae and planetary nebulae is supplemented by the fact that at the centres of both there are Wolf-Rayet stars and even their luminosity is identical. The only thing preventing us from assuming that the hundreds of planetary nebulae and their nuclei in the form of O and often Wolf-Rayet stars are the result of nova explosions, is that the masses of planetary nebulae are a hundred times greater than the masses of the ejected envelopes of these stars and they expand more slowly than the atmospheres of the latter. Here we have some mystery, but the similarity between these two forms of celestial bodies is too great for it to be random.

Why do novae cast their envelopes, can this happen to any star, can this happen to the Sun and are there changes in the configuration and core of a star after it cast off its envelopes?

All these questions arise because of lack of knowledge of the precise spectrum of a nova before its outburst. If we only knew in advance which of the faint stars was going to explode as a nova! We would at once obtain its spectrum and photograph it repeatedly, but it is impossible to do this for hundreds

of thousands of faint stars. Not knowing the spectrum of a nova before its explosion we can say nothing about its physical state before the catastrophe.

Nevertheless, there are certain data that make it possible to draw certain conclusions. It has been found that both before and after an explosion many novae irregularly change in brightness within small limits. The Sun does not behave in this way. There are stars which the author has called recurrent novae. Their outbursts recur irregularly after the lapse of several decades. The author has pointed out the identity of their spectra before and after explosions. There is nothing preventing us from postulating that the same thing occurs in other novae. Novalike stars differ from them only in the scale of the phenomena; otherwise they are completely the same.

The Moscow astronomers Parenago and Kukarkin have found that in semiregular variables the mean time between explosions is the longer, the greater is the change in their brightness. This relation is also satisfied by novalike stars. In particular, the repeated outburst of a star in the Corona Borealis, which they had predicted, actually occurred in 1946. If this relationship applies to novae, we can expect that they will experience repeated explosions approximately after 3,000 years.

Repetition of explosions with this frequency for certain stars can fully account for the observed annual frequency of explosions of stars in the Galaxy. This fact and also the assumed spectrum of novae before an explosion (it would be more correct to say between explosions), characterizing them as very hot stars, *excludes the possibility of the Sun's explosion.*

Walker in the USA recently made an interesting discovery. He found that the nova, that exploded in Hercules in 1934, consists of two almost identical stars with masses less than the solar mass.

But one of them, which explodes as a nova, is hot, and the other is a red cool dwarf. The stars of the pair are very close to each other. Therefore, because they attract each other strongly, they have unstable atmospheres. These latter evolve gas and its flows circulate round the stars, thus producing a common atmosphere as well. All novae, it seems, are close binaries.

Fascinating evidence came from the electrophotometer. Walker's findings suggested that all the "old novae" (i.e. novae after the explosion) are "feverish". The brightness of these hot stars fluctuates. Minor fluctuations with a period of several minutes seem to be superimposed on stronger, but slower, fluctuations.

A remarkable break-through came in 1966, it was associated with one of the strongest X-ray sources called Scorpius X-1. It was thought of as a tiny neutron star, which in theory is a "final" stage of stellar evolution when a star has exhausted its energy reserves. And it came as quite a surprise when in lieu of this source was found a fairly bright (about 13^m) star that resembled old novae in terms of its oscillations of brightness, and of colour and spectrum. Among the old novae, however, we as yet know of not a single intensive X-ray source, although some of them are not farther away from us and are as bright.

In addition, recurrent novae of Hercules type are close binaries, their satellites being cool as well. It may well be that the close binary character of the novae is somehow related to their explosions.

The whole body of evidence concerning novae and novalike stars denies the possibility that the cause of an explosion could be a collision of stars with one another or the falling of a planet on a star. If the collisions occur in the Galaxy they are so rare.

The factor responsible for the explosions of novae and novalike stars should lie within the stars themselves and the theory of the internal structure of stars leads to the conclusion that under certain conditions in the development of a star it is possible for an unstable state to set in. The "overproduction" of energy in its interior can then lead to a separation of the outer layers of the star.

The author believes that explosions occur in stars with a rather high temperature and a mean luminosity that makes it impossible to assign them to either the white stars of the main sequence or to white dwarfs. In the process of evolution only a small number of stars passes through such a stage or the stage is a very short one. That is why we almost never find them on the spectrum-luminosity diagram (Hertzsprung-Russel diagram). These stars are unstable and from time to time cast off their outer layers until the star finally attains a state of stability. After each explosion the star contracts somewhat and eventually, having contracted and cooled somewhat, it passes to the stage of a white dwarf. There can also be other methods of "production" of white dwarfs, however. The star is in

a Wolf-Rayet stage only between explosions, while they recur, and then, probably, not always.

In 1939 the German astrophysicist Biermann concluded that the stars become unstable when they run out of hydrogen in the process of its transfromation into helium, a process serving as the energy source in ordinary stars. When the energy output in a star becomes inadequate, the radiation of the star dwindles, i. e. its brightness decreases, and therefore it moves on the spectrum-luminosity diagram into a region intermediate between the white giants and dwarfs, where it experiences explosions passing to the stage of a white dwarf.

In 1945 the autor discovered that white and blue stars form a special sequence on the spectrum-luminosity diagram. The hottest stars form a continuous series on the diagram in the order of decreasing luminosity. The series begins with the most massive of the known stars, passes to the less massive and less bright Wolf-Rayet stars, and then to recurrent novae and typical novae, whose masses have not yet been established, and ends with the white dwarfs, the densest stars, which expend their energy the most sparingly.

We can interpret this discovery as suggesting that certain of the brightest and hottest stars are unstable. The light pressure in them is excessive and possibly nonuniform and the output of energy in their interiors leads to the escape of gas from their atmospheres, thereby transforming them into Wolf-Rayet stars. For the time being we cannot say under what conditions this change occurs and whether it is experienced by all hot giants. To be sure, mass is lost in this process and this is accompanied by a decrease in stellar luminosity. Eventually the star becomes more stable. Its equilibrium is disrupted with less frequency, but with greater and greater intensity with the passage of time, when somewhere within there is a gradual accumulation of conditions leading to instability. We have a new star with its infrequent explosions. After each explosion the star contracts and becomes denser until it passes into the stage of a white dwarf, so stable that explosions already cease completely.

Supernova Explosions

If we gathered together the best writers of science fiction and fairy-tales from all over the world and asked them to dream up something completely impossible, it is unlikely that any one of them could come up with anything as improbable as what we are about to describe. But this is not a fiction and not a fairy-tale but something that is occurring before our very eyes. Nature demonstrates "miracles" around us constantly and we are witnesses, although often we do not understand, or only understand much later, what miracle of nature we have seen.

Readers of books on astronomy have possibly grown accustomed to seemingly improbable distances of thousands of light-years to planets dissimilar from the Earth, to groups of coloured suns, thousands of times brighter than our own. But not only they but even experienced professional astronomers are overwhelmed when they ponder on the matter we are now going to discuss. The improbability and yet the certainty of these phenomena have become clear to us only recently, but many people from ancient times were witnesses. The Chinese chronicler Ming Tuan-Lin was one. On 4 July 1054 he wrote: "In the first year of the Ch'ih-Huo, in the fifth moon, on the day of Ch'ih-Ch'u a guest-star appeared to the south-east of the star Tien-Kuang and disappeared more than a year later." A countryman of Ming Tuan-Lin wrote: "It was visible during the day-time like Venus; light rays emanated from it in all directions and its colour was reddish-white. It was so visible for 23 days." Similar scanty records were made by Japanese annalists and Arabian eye-witnesses. Those records were found and read in 1942.

Many similar records (although not about such bright guest-stars), i. e. apparently about novae, have been found in old manuscripts. But almost a thousand years after the death of Ming Tuan-Lin astronomers studied in detail an unusual nebula visible in a telescope to the south-east of the Chinese star Tien-Kuang. We call it ζ Tauri and due to its singular configuration the nebula has been named the Crab Nebula by observers. This faint spot of light in the infinite blueness of the night sky shines like a crab of vague outlines in the bluish depth of the sea. Photographs of its centre reveal two stars of the 16th magnitude, i. e. 10,000 times fainter than stars scarcely visible with the naked eye in a dark, moonless night.

The Crab Nebula differs from ordinary nebulae, of which there are tens of thousands visible in the sky, in two ways. First, a comparison of its photographs,

A picture of the Crab Nebula taken in the blue part of the spectrum.

taken at an interval of 30 years, made it possible in 1942 to confirm an earlier discovered fact: the nebula is expanding appreciably in all directions from its centre, occupied by the two stars. Second, the nebula is unusual in that against the background of a bright continuous spectrum it is possible to see broad and double bright lines of chemical elements, among which, unlike other gaseous nebulae, hydrogen is scarce. The form of the spectral lines indicates that the nebula expands with a velocity of 1,300 kilometres per second, a velocity a hundred times greater than in other gaseous nebulae, which are also expanding.

By comparing the apparent angular velocity of expansion of the nebula with its linear velocity, determined from the spectrum, we can find the distance to the nebula (5,000 light-years), and then the luminosity of the two stars at its centre. The nebula is enormous and it takes 6 years for light to travel from one side to the other, whereas it crosses the diameter of the orbit of Pluto in the solar system in 11 hours.

Knowing the velocity of the apparent angular expansion of the nebula it is possible to compute when all its matter was concentrated in a single place, at the centre, where the two stars are visible. It has been found that this was about 800-900 years ago, i.e. approximately at the time when the Chinese saw their guest-star near this same place.

Can this be a simple coincidence? Is it possible that such an exceptional nebula randomly appeared at the same place and at the same time where the extraordinary nova appeared?

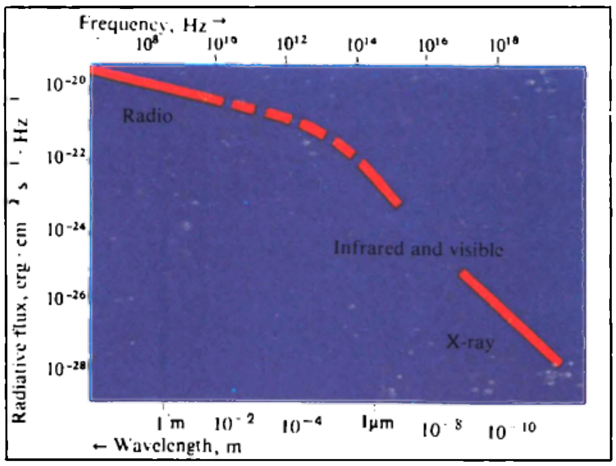
. After the explosion, this star left behind the Crab Nebula. In order for such a colossal nebula to be created there must have been a catastrophe far exceeding in intensity the explosions of ordinary novae.

Supernova, superstar–where do they come from, into what are they transformed and when do they disappear? These are questions to which there are no answers. If such a brightness can come from a sun similar to ours or if a supernova is transformed into a star like our Sun, after formerly having been something else, for the time being it is hopeless to attempt to see it. The stellar systems in which supernovae outbursts are observed are too distant for us to be able to detect in them a star of the Sun type. The Sun, if it was situated even in the nearest stellar system, would be several hundred times fainter than the faintest stars that we can distinguish at the present time.

If only could a supernova explode closer to us, in our Galaxy! Unfortunately, in the time during which astronomers have been interested in this phenomenon not a single such event has occurred. However, if it is assumed that the "guest-star" of 1054 was the factor responsible for the development of the Crab Nebula, and therefore was situated at the same distance from us, we find that its brightness was the same as that of supernovae. This was "our own" supernova.

The Crab Nebula has particularly strong emission of red rays, caused by certain nitrogen lines. This has suggested that we seek confirmation as to whether the

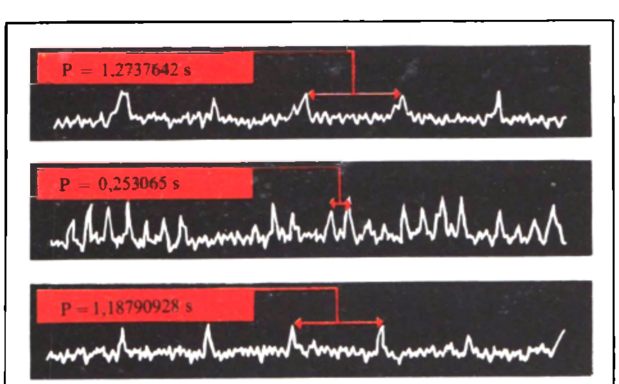
A picture of the Crab Nebula taken in the infrared part of the spectrum.



The energy distribution of the continuous spectrum of the Crab Nebula, which is seen to differ from the black-body law. The spectrum indicates that the radiation is due to nonthermal processes.

bright nova observed in 1604 in the constellation Ophiuchus was a supernova as well. In 1943 the neighbourhood of the constellation was photographed on the plates sensitive to red rays and on the photographs the investigators discovered an earlier invisible faint nebula. Its spectrum was similar to the Crab Nebula and its centre coincided with the place of the explosion of Kepler's supernova. At the centre of the nebula there are no stars brighter than magnitude 18 1/2.

The measured curves of some pulsars. Above: CP 0834; below: CP 0950; at the bottom: CP 1133.



A new star, brighter than Venus and visible even by day in 1572 in the constellation Cassiopeia, was also a supernova that exploded in our Galaxy and produced an expanding nebula. All the nebulae that resulted from supernova explosions are sources of powerful radio emission, notably the Crab Nebula. The radiation is produced by their proper magnetic fields braking fast electrons. A supernova explosion gives rise to many electrons in these nebulae travelling at nearly light velocity. This synchrotron radiation accounts for the continuous spectrum of the Crab Nebula. It is produced by the amorphous mass of this nebula, whereas the gaps produce gaseous radiation in bright spectral lines. The gaps permeate the amorphous bulk of the nebula, whose mass is about that of the Sun, i. e. is 10^4 times larger than that of envelopes ejected by conventional novae.

Strong radio emission is a telltale feature of expanding envelopes ejected by supernovae. From this indication a number of such envelopes were found of supernovae that had exploded too long ago, or had been overlooked for some reason or other. The most remarkable of them is the Cassiopeia A Nebula. It seems to be closer to us since it is the strongest radio emitter as observed from Earth.

That there exist supernovae was found earlier than stars in our stellar system were identified (in the 1940s) as such that exploded in 1054, 1572, 1604 and that had initially been believed to be conventional novae. The name supernova has first been applied to stars that exploded in some other stellar systems, not in the Galaxy. As early as the 1920s photographs showed explosions of stars in distant stellar systems that are comparable with our Galaxy in size and number of stars.

In other, nearby galaxies nova explosions have been observed, exactly as described in the previous section. For example, in the spiral galaxy M 31 about 30 novae explode annually. At their maximum they are brighter than all the other stars of the galaxy. But supernova explosions, which are significantly more rare occurrences, at their maximum are tens of thousands times brighter still. Their luminosity is for several days equivalent with the radiation of several milliard suns. At times their radiation exceeds that of the entire stellar system. However incredible this phenomenon seemed because of the intensity of the explosion, scientists had to recognize its reality – facts are

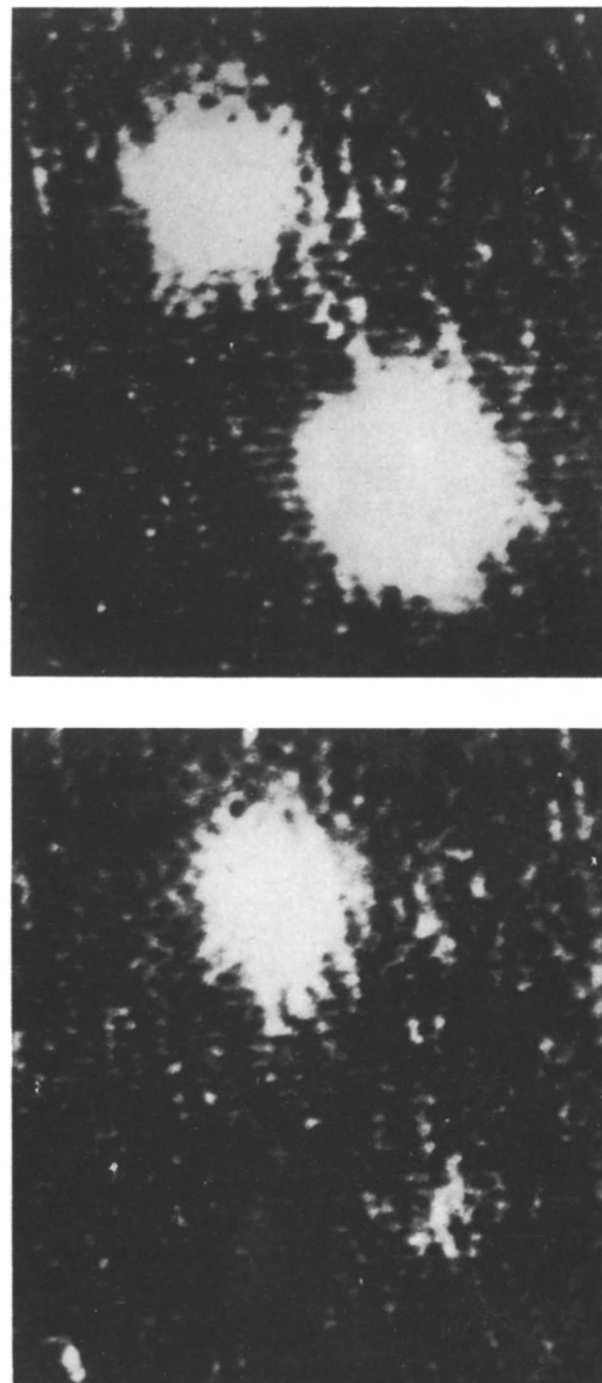
stubborn things. This phenomenon dwarfs grandiose catastrophes in normal novae.

Such phenomena, more incredible than the tales of Arabian Nights, reveals Science!

Supernovae (a misnomer!) explode exceedingly rarely; the average occurrence for one galaxy consisting of milliards of stars is about one explosion per 300-400 years, but in larger galaxies they occur more frequently than in small ones. This was found by Zwicky, who back in the 1930s began systematic hunt for such outbursts. On his suggestion in 1961 an international "supernova service" was initiated, which involves 11 member-countries. As a result, by 1980 about 500 supernovae were recorded, more than normal novae recorded in the Galaxy. One should make allowances for the fact that supernovae at their maximum are no brighter than the 13th magnitude, and the novae being discovered are normally much brighter.

Being exceedingly far away and faint even at their brightest supernovae do not lend themselves to a detailed study either. Much important evidence has been gleaned however. When correlated to the data of radio-emitting expanding nebulae (ejected by the supernovae of the Galaxy even before the telescope came in), this evidence makes it possible to study the phenomena in question as a whole.

Minkowski, who investigated the light curves and spectra of supernovae using a 5-metre telescope, has established two types of supernovae differing in light curves and emission spectra. Type I spectra do not exhibit distinct lines and have not been interpreted for a long time. Type II spectra are similar to those of normal novae and vary in the same way, only their bright lines are wider and suggest ejections of gases with velocities of several thousand kilometres per second. The supernovae of the Galaxy in Taurus, Cassiopeia and Ophiuchus seem to have been type I supernovae. They are somewhat brighter than type II supernovae, explode much more rarely, although in any part of a stellar system, whereas more frequent type II supernovae only explode near the flat layer of those galaxies where such a layer is available. Since the stellar population of flat layers of galaxies differs from



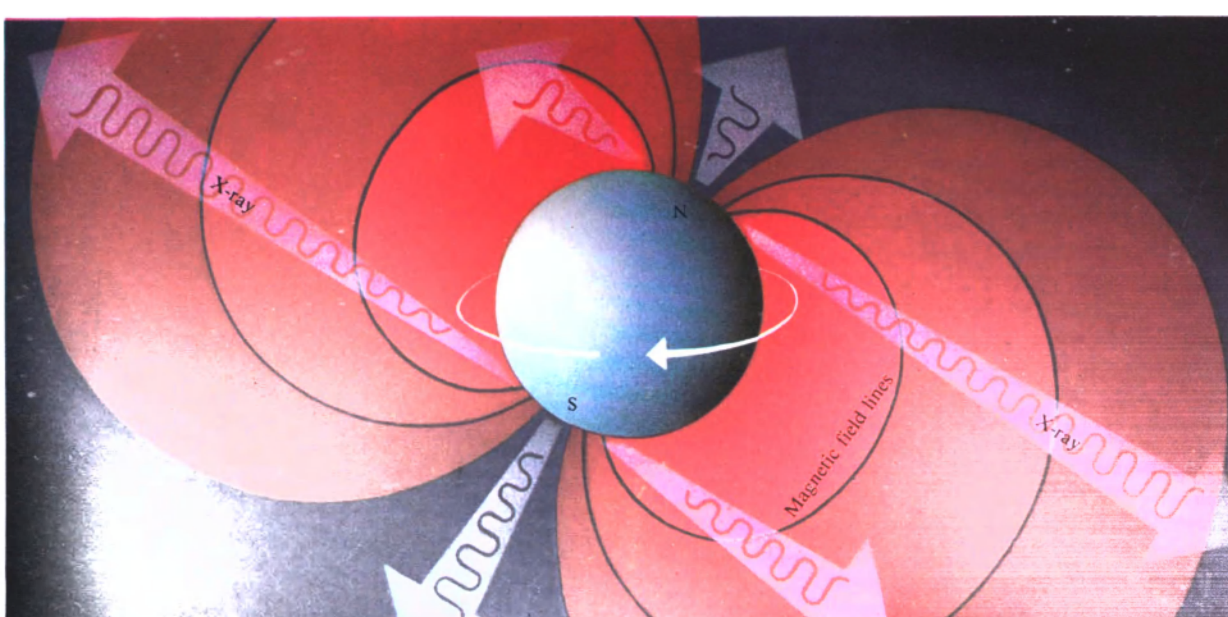
"The lighthouses of the Universe": pulsar NP 0532 (neutron star about 10 km in diameter) discovered in 1967 (in 1969 optically) in the Crab Nebula.

that of their spherical component, the stars that explode as type I and II supernovae are different. Zwicky concluded that there are perhaps five types of supernovae, not two, which adds to the complexity of the problem. A supernova eruption produces a prodigious energy of about 10^{50} erg.

The theoretically predicted radio emission produced by supernovae was first recorded in 1971. Since the advent of the telescope no supernova outbursts have been observed in our Galaxy. So far we have been observing them optically only in other unbelievably distant stellar systems, which are so far away that a sunlike star could not be seen through the most powerful telescope. The radio flux of 1971 came from a star previously found in the spiral stellar system designated M 101. The light from it takes several million years to come to us and it is impossible to see in it stars that are even a thousand times brighter than the Sun. However, photographs taken using a powerful telescope recorded the supernova near its maximal brightness. The radio flux from it at 21 centimetres was about 10^{-28} watt per square metre per hertz. It seems that the radiation is not of thermal nature.

Zwicky postulated that the liberation of energy in the form of heat and light, observed at the time of a supernova explosion, cannot be explained by commonly accepted sources of stellar energy. He assumed that in the particular case energy is liberated at the time of transformation of a star consisting mostly of atomic nuclei into a star consisting of neutrons. When all the hydrogen has been transformed into helium, in stars of a particular mass due to extraordinary densities and temperatures free electrons will, as it were, be forced into the nuclei of atoms under high pressure. The electrons, upon being forced into the nuclei, neutralize their charge and transform them into neutrons. The neutrons, having the size of atomic nuclei but having no electric charge, which would be an obstacle to their merging, can be led into contact with one another far more easily than the electrically charged nuclei of atoms. External pressure compresses the star very hard and immediately gravitational energy liberates violently. The radiation excess inside the star ejects the outer layers of it into space and the remainder of the star

Rotating neutron star.

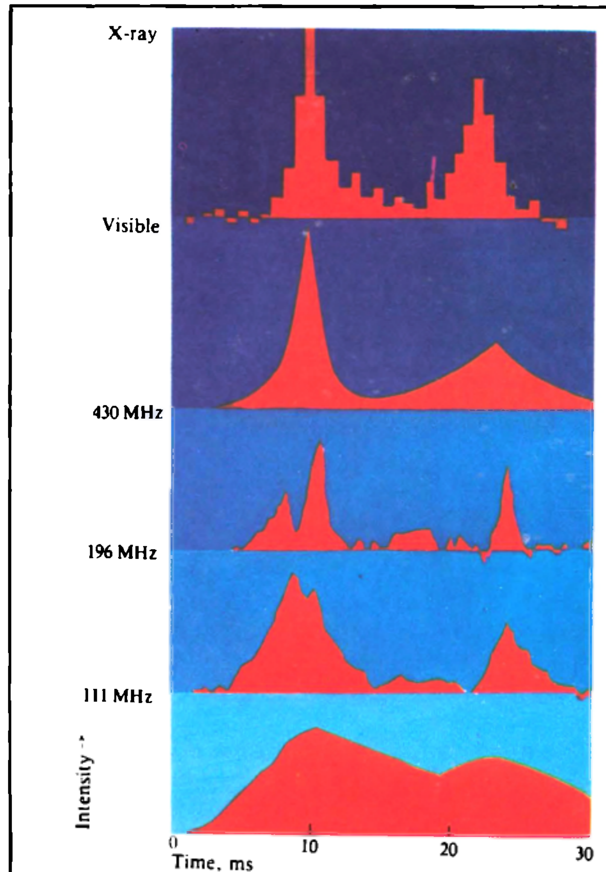


collapses toward the centre, like a house of cards, and contracts to the density of neutrons, i.e. 10^{14} grammes per cubic centimetre. In Zwicky's opinion the star's diameter decreases to about 10 kilometres!

This star would have the size of an asteroid. A thimbleful of neutrons would weigh a hundred million tonnes. The entire Earth's mass would under these conditions fit inside a sphere with a diameter of 150 metres. This is not far from the small bundle containing the entire Earth's gravity that the Russian epic hero Svyatogor attempted to lift.

It has been calculated that a neutron star should be emitting a powerful flux of X-rays, which would be the only indication of the whereabouts of this midget. In the mid-1960s several most powerful X-ray sources were found in the skies. One of them coincided with the above-mentioned nebula Cassiopeia A, and another with the Crab Nebula. But the X-ray emission of the internal regions of nebulae left by supernovae comes from the nebulae themselves and, it seems, has synchrotron nature as well.

Soon the discovery of radio pulsars, one of which happened to be inside the Crab Nebula (it "doubles" as an optical pulsar and X-ray pulsar), brilliantly confirmed the predictions of my friend late Fritz Zwicky. Neutron stars and then the so-called "black holes" turned out to be reality.

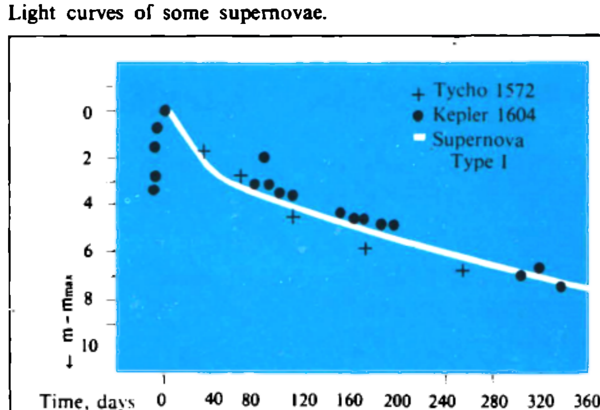


Average pulse profile.

Stellar Tops – Pulsars and Black Holes

The first pulsar was discovered by the English astronomer Hewish and co-workers at the end of 1967. This was a major, unexpected break-through. It was found that a seemingly "empty" patch of the sky sent out short radio pulses recurring with an exceptional stability with a period of less than one second. A first idea that would cross the discoverers' minds would be if these might be some artificial signals from some extraterrestrial civilization. Pulses of the known pulsars, which are now over one hundred in number, are very stable and vary widely in intensity and the degree of polarization. The periods vary from 4 to 0.02 second, i.e. they are fairly stable. However, periods normally increase, rarely decrease, sometimes in jumps. Significantly, apart from radio pulsations, X-ray pulsations have been found. Lastly, the pulsar designated NP 0532 showed an optical pulsation – after it has been identified with a faint star

Light curves of some supernovae.



that is a remnant of the supernova whose explosion gave rise to the Crab Nebula. But this is so far the only visible pulsar. Pulsars sometimes show abrupt changes in their periods. This may be a result of "stellar quakes" that lead to a quick restructuring of a neutron dwarf-star, or to a powerful plasma ejection. Pulsars noticeably concentrate about the Milky Way, therefore they are to be found inside the Galaxy. Their distances are only estimated roughly using indirect techniques. They are generally hundreds, more often thousands, of light-years away and the pulsars, save for one, are as yet invisible because their light emissions seem to be exceedingly weak. Most of authorities agree that pulsars are tiny neutron stars with a diameter of several kilometres that rotate with periods of a split second. They could be called stellar tops. A pulsar's mass must be comparable with that of the Sun, although the density at its centre must be about 10^{15} grammes per cubic centimetre and its magnetic field must be very strong (up to 10^{13} oersteds). The pulsed radiation must come from radioemitting regions on its surface that rotate with the whole of the star relative to us. The magnetic field must rotate too. Models of pulsars and mechanisms of their emissions are being studied theoretically.

General Relativity admits that under gravity, when the pressure of a gas becomes smaller than gravitational one, the compression occurs in a catastrophic manner – the star collapses. If the star's mass is more than 1.2 solar masses, but less than two solar masses, the star is compressed to become a neutron star. But if the star's mass is larger than mentioned above, the compression proceeds further past the so-called Schwarzschild's radius. Neither light nor matter can escape the star, it becomes invisible. Earlier in the book we referred to comets as "visible nothing", but here we have "invisible something", principally invisible, tiny but with enormous mass. This "something" cannot even be dubbed a star, and so this initially purely theoretical entity began to be called "black hole" (not to be confused with "black body" in physics!) because it does not glow and is a "special point" in space. Discussions of possible properties of black holes and search for them are now popular among physicists and astronomers.

It has been assumed that when some gas falls into a black hole it can radiate sufficient energy for the black hole to be visible, and that a black hole revolving with a short period in a pair with a visible star can be found from its motion. It has also been found that some eclipsing variable stars produce X-ray radiation like pulsars. It is also assumed that around a normal star a black hole is orbiting at a close distance. It radiates just like a pulsar produced from a star of some unknown type that has exploded like a supernova. The X-ray emitter here is the disk of hot gas – plasma – that flows from the visible star to the black hole. Masses of objects suspected of being black holes (they enter the pairs) are determined from revolution periods and velocities – the quantities that are dependent on the masses of the component bodies. The masses of black holes have been estimated to be several solar masses. In 1974 a radio pulsar was found with the frequency of radiation varying due to its orbital motion. Analysis of its motion showed that its mass is about that of the Sun. This agrees with the concept that black holes are collapsed stars, as it were, dead stars – the last stage of their life.

It was reported at the International Astronomical Congress in 1973 that in the stars with an initial mass of 10 solar masses nuclear fusion in their cores leads to the formation of iron cores inside them with a density of 10^8 grammes per cubic centimetre. As the temperature rises above $5 \cdot 10^9$ K there occurs a gravitational collapse in the form of supernova explosion and neutron pulsar formation, if the mass of the iron core is less than two solar masses, and black hole formation at a larger mass. No observational evidence is as yet available to suggest what the stars are like just before they erupt as supernovae.

4. The World of Star Clusters and Diffuse Gases

Open and Globular Star Clusters and Associations

Open and globular stellar clusters differ from one another about as disorderly crowds differ from ranks of soldiers.

Open clusters are to be found within the Galaxy where they alternate with single stars, they are sort of major cities in a country. It is for their position in space that they are sometimes called galactic, and for the low concentration of stars to the centre of the cluster they are called open. The stars in them run into thousands and they are scattered in space disorderly, just like the tents in a Gipsy camp.

An example of open clusters is the Pleiades. In autumn they rise in the evening and in winter they are high in the sky in the evening. An unaided eye of average sight sees in this cluster six stars, and a keen eye, from seven to eleven. But hundreds of stars of different magnitude flicker in the field of vision of the telescope. That a star belongs to a given cluster is found from the general nature of its motion in space. So we can single out stars, more near or distant, that happen to project on the stellar cluster.

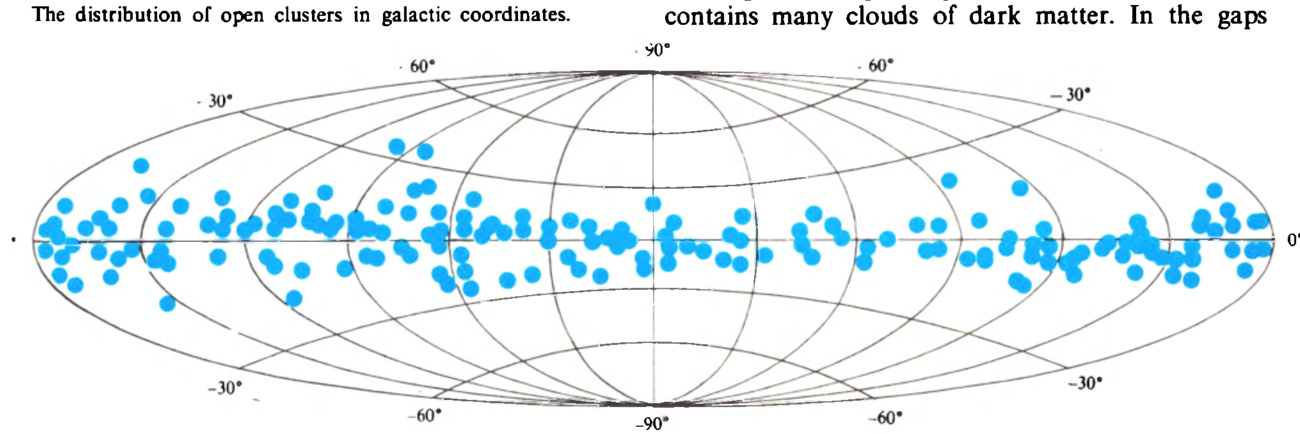
By measuring the magnitudes and spectra of stars in clusters or by determining colours, which is simpler, some sort of a Hertzsprung-Russel diagram can be worked out for them. More often than not it is similar to a diagram plotted for the near-solar space. This diagram generally appears to be incomplete since it lacks the giant branch (and, of course, since it is impossible to see white dwarfs in distant clusters).

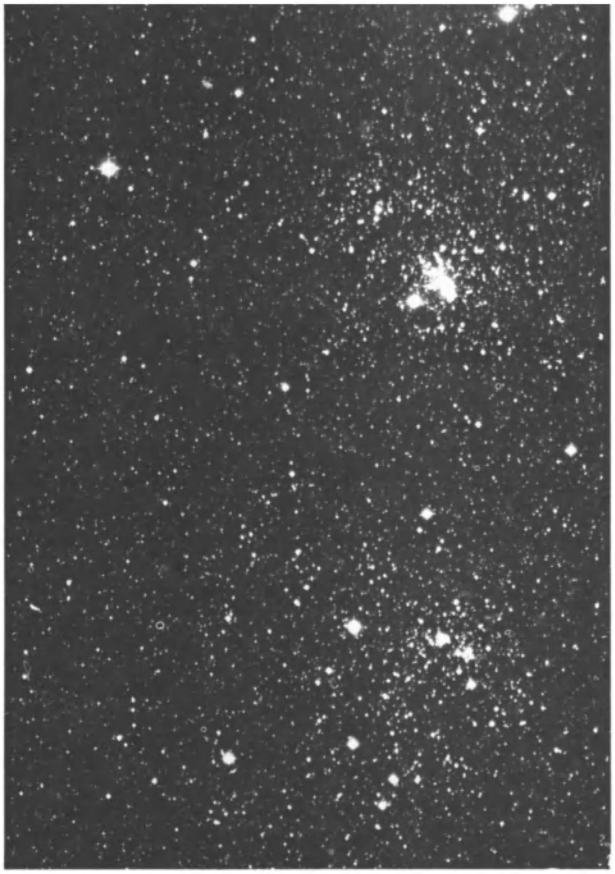
Comparing it with a diagram for the solar neighbourhood and, so to speak, equating them we can determine the difference $m - M$, i.e. the difference between the apparent stellar magnitude of stars of each spectral class in a cluster and their absolute stellar magnitude. Using this difference, as we have already seen, we can easily work out the distance to the cluster. Further, if we know the distance and measure the apparent angular diameter of the cluster we can work out the linear diameter of the cluster in light-years. For example, the Pleiades are separated from us by 320 light-years and the diameter of this star group is about 30 light-years.

Around the red Aldebaran, the brightest star in the constellation Taurus, we can easily see the few (and more scattered than the Pleiades) stars of the cluster Hyades. All told, we know more than 1,000 galactic clusters, but we still do not know a multitude of those that are more distant or faint, or blotted out by dark nebulae.

Ambartsumyan has distinguished groups of stars in the sky that he called associations. The stars in an association have the same physical features and are more scattered than those in galactic clusters. The latter themselves frequently enter an association. Ambartsumyan called groups of hot stars containing O stars or early B stars O-associations, and groups containing variable T Tauri stars T-associations. Associations are delineated in the skies from the apparent clustering of such small groups of stars. Such visible clustering of O and B stars is in need of thorough checking. The point is that the Milky Way contains many clouds of dark matter. In the gaps

The distribution of open clusters in galactic coordinates.





Open clusters in Perseus (h and χ Persei).

between them are "visibility corridors". Through a corridor we can see distant hot stars among nearer ones, and so the apparent stellar density in a given region appears to be high, whereas the stellar space density is not that high and there is no such dense clustering.

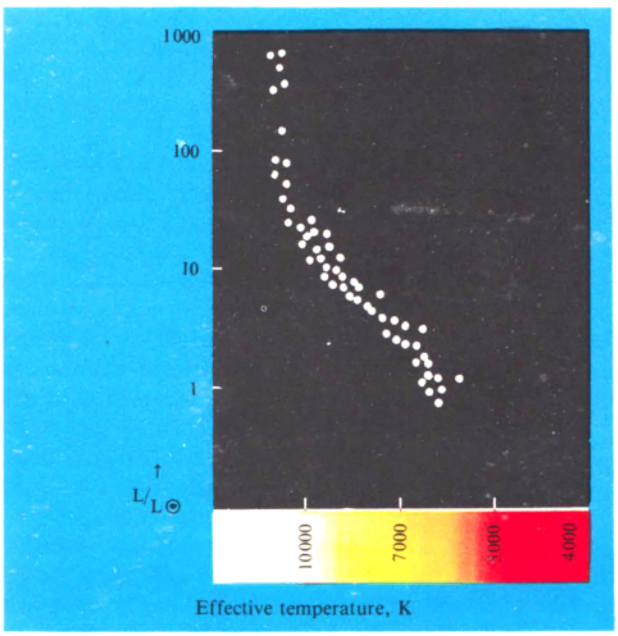
But even if it does occur, the mutual attraction between the stars in an association is small, since they are widely separated and so the stars will gradually drift away from this region. It is as yet difficult to establish this recession and it continues to be a matter of debate. According to Ambartsumyan, O-associations give rise both to hot stars and to cooler stars. Apart from stellar clusters, these associations are the cradles of stars. The associations are intermediate in size between galactic clusters and larger star clouds.

Globular clusters, of which about one hundred are known, have a typical representative – the cluster in Hercules, which is seen through the binoculars as a fuzzy star of about 6th magnitude. Only a strong telescope, and especially photographs, show that there is an entire cluster here in the form of a sphere, the density of stars increasing towards the centre. Hundreds of thousands of stars are here, of which we only see the brightest. Fainter stars, in particular such as the Sun, are invisible. Owing to enormous distances and number, especially near the centre, the stars merge into a continuous bright glow.

Distances to the globular clusters have long been an enigma until Cepheid variables have been found among them. Just imagine that in a tiny region of the sky occupied by a cluster you find one, two, three, etc., Cepheids, whereas in the vast expanses around the cluster you find none. Could this be a chance coincidence?

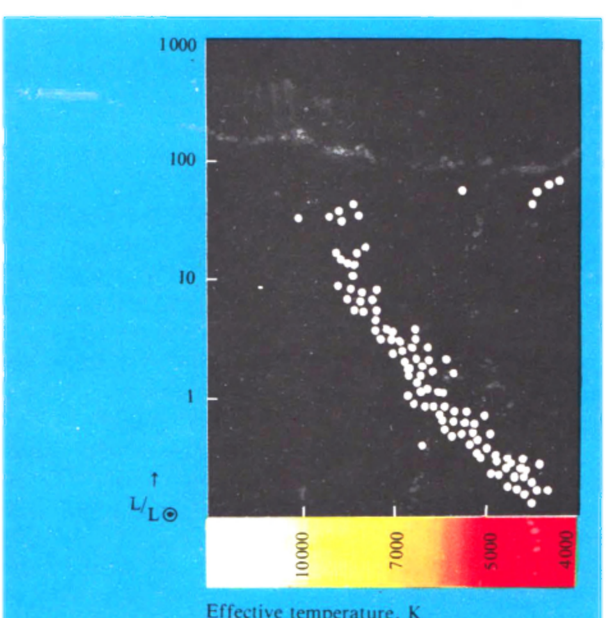
One Cepheid variable that is nearer to us than the cluster or more distant may project on the cluster – this will be a random occurrence. If of all the Cepheids in this region a second one also projects on the same patch of the sky, this can be called a coincidence. But if scores of them project there, this cannot be, as it were, a habit, because stars have none. This can only mean that Cepheids actually lie within the globular cluster itself and are its members. The presence of Cepheid variables makes it possible to determine distances to a number of globular clusters, and then their sizes. The distances to those that contain no Cepheids can, according to Shapley, be determined from the apparent brightness of the brightest stars. The distances to the most distant clusters, which appear as spots in the sky and where no individual stars can be

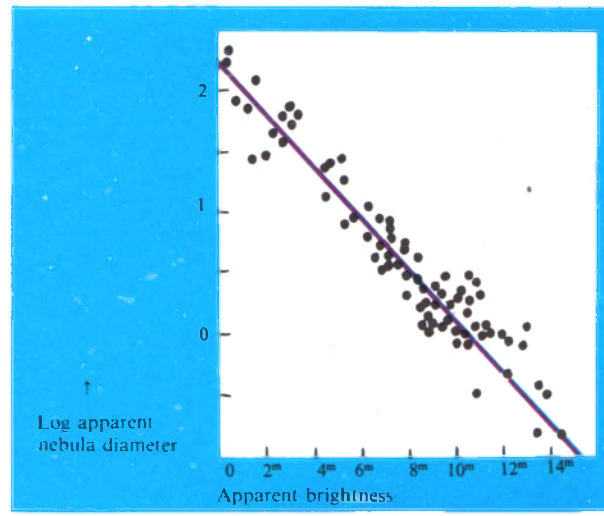
M 92 on 20/21 September 1966.



The Hertzsprung-Russel diagram of Pleiades.

The Hertzsprung-Russel diagram of Praesepe.





The variation of apparent size of a gaseous nebula and apparent brightness of the illuminating star.

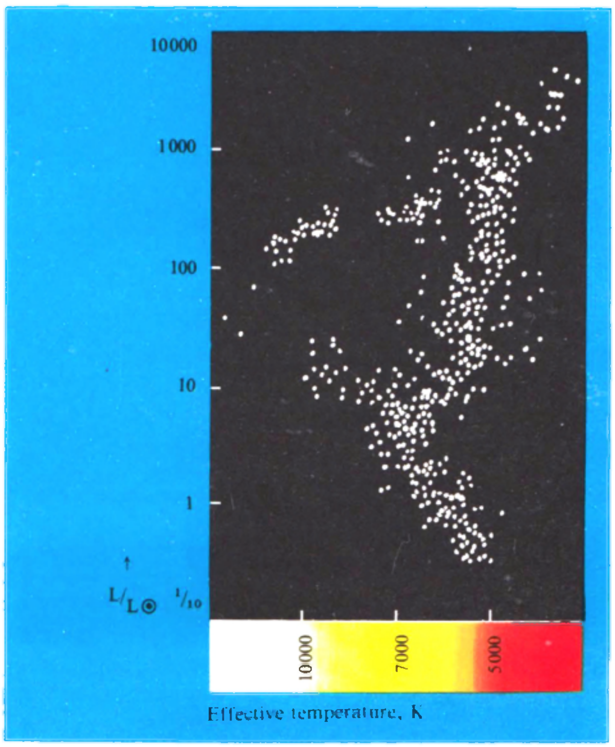
distinguished, can be determined from the apparent angular size and combined apparent brightness, since true linear dimensions and total brightness are about the same with all the globular clusters.

One of the nearest globular clusters, the one in Hercules, is separated from us by 20,000 light-years, its diameter being a hundred light-years. The most distant star clusters lie 230,000 light-years away from us.

The Hertzsprung-Russel diagrams of stars in globular clusters differ markedly from galactic clusters and the solar neighbourhood. Also, some different types of stars are present there. Comparison of appropriate diagrams for different stellar systems enables important conclusions to be reached as to their life cycles. This will be discussed later in the book.

The German-born American astronomer Baade was the first to assume the existence of two types of stellar population differing in their spatial distribution. The studies of the Soviet astronomer Kukarkin and co-workers indicated that objects with different physical behaviour (for example, variables of different types, clusters of different types, etc.) enter more diversified components of the Galaxy: flat, spherical and intermediate ones. There are good reasons to suppose that the objects that comprise the different components have different origins and age.

So globular clusters and short-period Cepheids are



The Hertzsprung-Russel diagram of the globular stellar cluster M 3.

included into the spherical component filling the space within the sphere with the centre at the centre of the Galaxy. Other celestial bodies, for example hot giants, dust and gaseous nebulae, are part of the flat component, concentrating mainly in the thin layer along the plane of the Galaxy. The distances of the globular clusters suggest that they concentrate towards the Galaxy's centre, but reach as far as the boundaries of our stellar system beyond which star-free space lies. The size of a system of globular clusters thus defines the overall dimensions of the Galaxy, of the densely populated stellar home of ours.

What then lies beyond? Are there any other stellar homes, other galaxies, are they like ours or not? We shall discuss this in the next chapter.

Clearer About the Nebulous World

Nebula is the term assigned by astronomers to anything that is nebulous in the sky, regardless of what it consists of, provided it occupies a fixed position in the sky among the stars. As the observational techniques improved, certain (actually most) of the visible nebulae have been found to be distant stellar systems, similar to our Galaxy.

But in certain bright nebulae near or within the Milky Way itself astronomers have found spectra consisting of narrow bright lines against a dark background.

If we are to define their physical nature, not only their appearance, we will have to call them gaseous nebulae. Since they lie within our stellar system, the Galaxy, they are referred to as galactic nebulae to distinguish them from extragalactic nebulae that are essentially giant stellar systems like ours, and commonly themselves contain gaseous and dusty nebulae.

Thus, gaseous nebulae, as it is best to call them in order to avoid misunderstanding, consist of thin gas. According to their configuration they are broadly divided into irregular or *diffuse*, and *planetary*, which are circular, oblong or ring-shaped with a nucleus star at the centre. It was highly inapt to call them planetary nebulae because they have no relationship whatsoever to planets.

To clarify the picture we will say that gaseous nebulae mainly consist of ionized hydrogen. Extensive regions of hot ionized hydrogen are called H II (H is the chemical symbol of hydrogen) and the regions filled with cool, nonionized, invisible hydrogen are called H I.

Nebular Gas

Spectral lines show that gaseous nebulae consist of hydrogen, helium, nitrogen, oxygen, carbon, and certain other chemical elements. But the brightest spectral lines are two green lines that for more than half a century have been attributed to an unknown gas, since not one chemical element in the laboratory has revealed these lines under any conditions. As the unknown gas on the Sun was called "solar" or "helium" so the unknown gas of nebulae was called "nebular" or "nebulium".

The years passed but the mystery of nebulium was not solved. Only advances in spectral theory and atomic theory made it possible to unmask the stranger.

Much like coronium on the Sun, the "nebular gas" was dressed up in the unfamiliar form of green spectral

lines and was concealed by them. It became clear long ago that the nebulium was masking something familiar, since there was no place for it in the periodic system of elements.

The mask was torn away in 1927 by Bowen, who computed the wavelengths of almost all the spectral lines that could be associated with the chemical elements already discovered in nebulae. It was found that the green lines are the "forbidden" lines of doubly ionized oxygen. In nebulae doubly ionized oxygen emits both "permitted" and "forbidden" lines, and the latter are even brighter than the former. In the unmasking of solar "coronium" it was found that such forbidden lines can be observed in laboratories only with the greatest difficulty. For emission of these lines the gas should be extremely rarefied and the energy incident on it from the star also should be extremely rarefied, i. e. the gas should be situated quite far from the star, where illumination is very weak. Up to the present time it has not been possible reliably to induce the appearance of the green lines of doubly ionized oxygen, because on Earth we still cannot simulate the necessary conditions. Our best air pumps are still unable to create even approximately the rarefaction of gas which exists in gaseous nebulae. Moreover, this gas, which from the terrestrial point of view is far more rarefied than that which we call a vacuum under the bell of an air pump, shines brightly. We see it at distances of thousands of light-years, and if we refer to gaseous nebulae discovered in other galaxies – at distances of millions of light-years.

In the laboratory we have not yet observed numerous other spectral lines of nebulae, less bright than the green lines, and also ascribed to the same nebulium. They have been found to be the forbidden lines of the same oxygen, but singly ionized or neutral, and also the forbidden lines of other known chemical elements. Gaseous envelopes ejected by novae give exactly the same spectrum as gaseous nebulae and sometimes in the spectra of novae all the brightest lines, and in certain cases all the visible lines, belong to the category of forbidden lines.

At present all the spectral lines of gaseous nebulae have been identified. There are more than one hundred known lines. From these lines we determine qualitatively the chemical composition of the nebulae. It is primarily characterized by light elements, but, just as with stellar atmospheres, nebulae can also contain some other chemical elements, although their lines are not observed in the spectrum. This is because the lines are either weak or belong to a band that cannot be studied under the terrestrial conditions: in the ultraviolet (which is absorbed by the Earth's atmosphere) or infrared (which has strong absorption lines due to atmospheric vapours).

It is much more difficult quantitatively to determine the chemical composition of nebulae, i. e. percentages of various chemical elements. Other things being equal, the brighter, or more intense, are the lines of a given species of ions, the more there are of these ions, since each quantum of light in the line is produced by the radiation of one ion. But the line is produced in a wide variety of physical conditions, and many ions give no lines in the range observed. The total number of atoms of a given element equals the sum total of all the neutral and all the ionized atoms of the elements.

It can be assumed that within the accuracy of calculations there is no significant difference between quantitative chemical compositions of nebulae and stars. It would be of especial interest to compare the chemical compositions of nuclei of nebulae and their envelopes, since the envelope matter (considering its expansion) has undoubtedly separated from a star some time ago. It is all the more interesting because among nuclei with the Wolf-Rayet spectrum some contain carbon without nitrogen, and others contain both carbon and nitrogen, nitrogen even dominating in one case. Unfortunately, such comparisons of chemical compositions are not easy, in particular because the spectral lines of a nebula are superimposed on the nucleus's lines, which are few as it is, and there is no way of separating them out. It is well known that solar prominences are unusually rich in ionized calcium in comparison with its content in the chromosphere from where they are ejected. Prominences come in hydrogen and metallic versions. This sort of a difference is also possible in planetary nebulae.

Luminescence and Nature of Gaseous Nebulae

The emission spectrum of gaseous nebulae and the fact that their brightness is greater than that of adjacent stars, which could be suspected to cause

their luminescence, refutes the possibility of their luminescence by reflected light. However, the rarefaction of gas suggested by the spectrum excludes its being hot and fully self-luminescent. The Americans Hubble, Bowen and Menzel, the Dutch scientist Zanstra and the Soviet astronomer Ambartsumyan have established the principal characteristics of the luminescence and the nature of gaseous nebulae.

Gaseous nebulae shine to a certain degree in the same way as comets or the gas in a fluorescent lamp. Their luminescence is forced.

Stars force the gases to luminesce: in planetary nebulae – the stars situated at their centres, and in diffuse nebulae – stars situated somewhere within them or even in the neighbourhood. But without question such a star should be very hot. And that is the case. The stars exciting the luminescence of gaseous nebulae are of spectral class O or B0, not later, i.e. their temperature is above 25-30 thousand degrees. At such high temperatures the energy maximum in the spectra of these stars lies in the ultraviolet region, invisible to the eye. The nebula absorbs powerful fluxes of invisible ultraviolet rays and its atoms then radiate the absorbed energy in the region of rays visible to the eye, for example, green lines. The minimum portion, or quantum, of visible rays contains less energy than a quantum of ultraviolet rays. Therefore, by the law of conservation of energy the nebula should emit *more* quanta than it received. The perception of brightness by the eye is dependent on the number of quanta incident on it per second. That is why gaseous nebulae in visible rays shine more brightly than the stars causing this luminescence. However, the energy of visible rays radiated by the nebula is equal to the energy of the ultraviolet rays absorbed by the nebula.

Under the extremely high temperatures of the star the gases of the nebulae are ionized very strongly. For example, it is possible to observe quadruply ionized oxygen. Hydrogen shines when its ions capture free electrons which are passing by. But the forbidden lines of oxygen are emitted after the atoms or ions of oxygen are excited by the energy of collision with slowly passing free electrons. In order to emit the green line of "nebulium", an oxygen ion should be in an excited state, and hot for no less than a few minutes. During this period, therefore, it should not be "alarmed" by impacts of light quanta, other atoms or electrons. In order for collisions to be so rare the quantity of particles in a unit volume (i.e. the gas density) should be very small. Computations have shown that the density of gas nebulae is 10^{-19}-10^{-22} grammes per cubic centimetre. At this density, hours will pass from one atomic collision to a next. As a result of the remoteness from the star, its light quanta also fly far from one another and rarely collide with ions. Thus, atoms have all the conditions necessary for their emission of forbidden lines, i.e. forbidden on Earth where there is a high density of gases. Under terrestrial conditions atoms collide far more frequently than people in a market-place, and in a nebula, by comparison, they collide more rarely than wandering musicians meet with one another. In the air molecules travel a path of millionths of a centimetre from collision to collision, but in a nebula the so-called "free path" is measured in millions of kilometres.

As we have already mentioned, the mass of the colossal volume of gas forming a planetary nebula, as a result of its rarefaction, constitutes only one hundredth of the solar mass. The masses of large diffuse nebulae can be hundreds of times greater than this. Dokuchayeva, applying the theory of Ambartsumyan, has estimated, for example, that the mass of the Orion Nebula is 500 solar masses.

Very frequently, diffuse gaseous nebulae are confused with dust nebulae, which shine by reflected light, and even with dark dust nebulae. Does the gas condense into meteoric dust anywhere? Other considerations lead us to this thought.

Apart from bright lines some planetary nebulae emit a faint continuous spectrum. It might appear incredible that these nebulae, being highly transparent, could contain noticeable amounts of dust reflecting the light coming from the nucleus. The puzzle was solved in 1950 by the Estonian physicist Kipper.

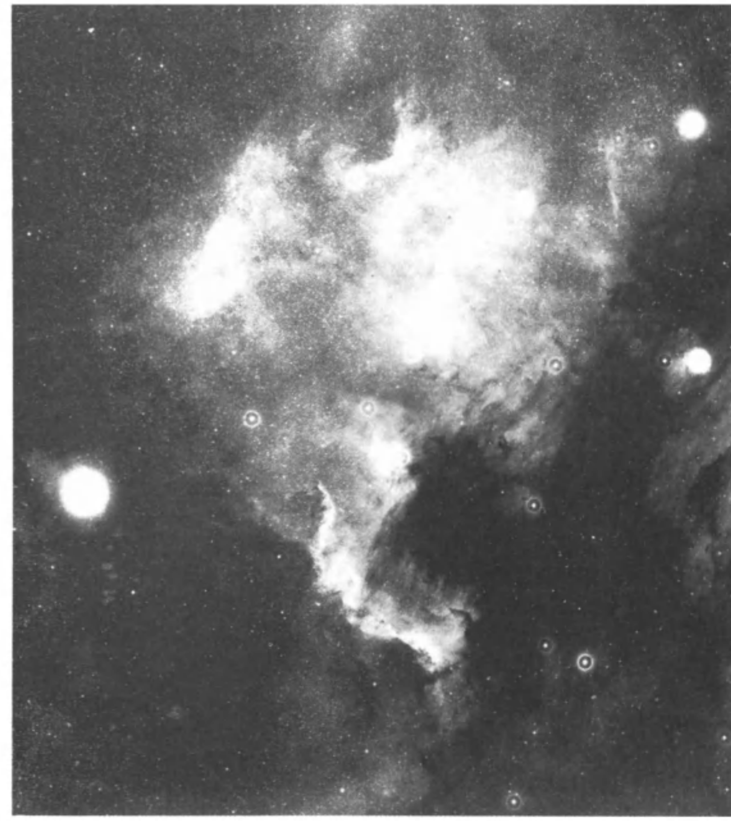
It has earlier been known that some atoms may undergo a forbidden transition, radiating two quanta rather than one, so that the total energy of these quanta is equal to the energy difference between the two respective levels in the atom. The probability of the two-quanta transition is small, but not

infinitesimal. The lifetime for one of sublevels of the second state in the hydrogen atom is 0.12 second. Transition from it to the ground state gives two-quanta radiation, but in different cases the energy is unevenly distributed between the two quanta. In this way a multitude of atoms radiate a wide variety of quanta of different frequencies so producing a continuous spectrum. To a lesser degree it can be produced by helium atoms too, both neutral and ionized. The brightness of the hydrogen continuous spectrum varies with the population of the second level, which in turn is proportional to the frequency of recombinations of protons, hence to the brightness of hydrogen lines. Quantitatively, this theory also agrees with observations, especially if certain subtle features of the process are taken into account along with the fact that hydrogen atoms recombine radiating a faint continuous spectrum (it is also produced by electrons decelerated by flying near atoms that are, however, unable to capture them). So the puzzle of the continuous spectrum in the gaseous nebulae has also been solved by the theory.

Diffuse Gaseous Nebulae

Normally fairly fragmentary, gaseous diffuse nebulae concentrate near the galactic equator. They come in a wide variety of sizes and shapes. The best known of them are the nebulae Orion, Lagoon, Omega, Trifid, Pelican, North America. But there also exist more clearly delineated objects in which brightness grows towards the periphery (peripheric nebulae), such as the Rosette Nebula. In its centre lies a galactic cluster consisting of O and B hot stars. Also, there are a few filamentary nebulae. The best known of them is NGC 6960 and 6992 or Network in Cygnus which is, however, believed to be a remnant of a supernova.

Photographs taken through a red filter suppress the night-glow emission and enable very faint nebulae to be revealed at the red hydrogen line. Many of them have been found by Shain and co-workers at the Crimean Observatory. Together with Fesenkov and Rozhkovsky, Shain issued beautiful atlases of photographs of these objects that show fine details and allow distinct indications of turbulences to be perceived.



North America Nebula, NGC 7000.

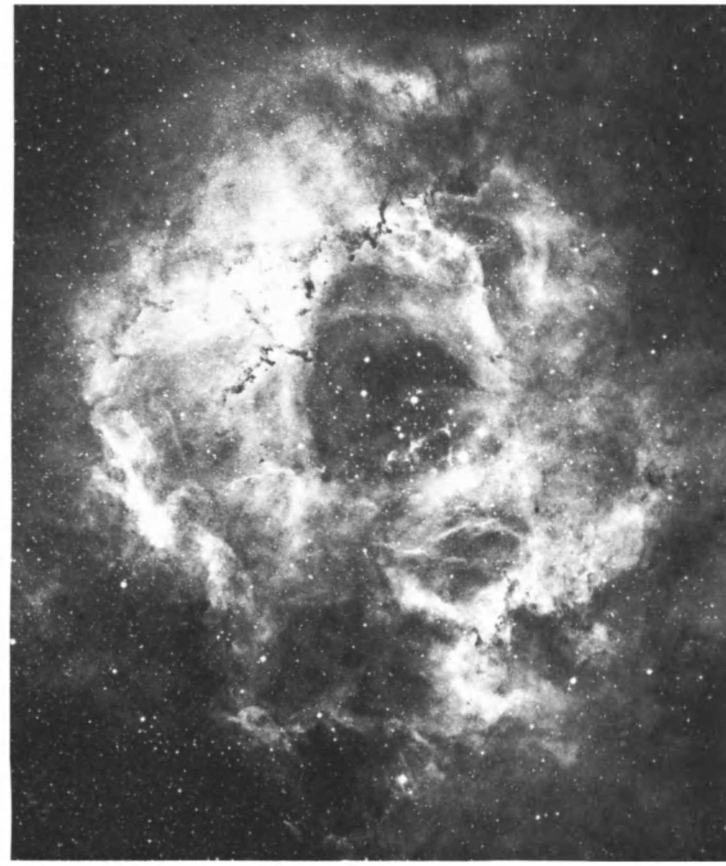
In the Orion Nebula such turbulent motions also manifest themselves in that radial velocities differ from place to place.

About 300 different diffuse gaseous nebulae are known, but their number and sizes available in catalogues are fairly arbitrary because there often occur nebula complexes, each of which can be thought of as an individual nebula. On the other hand, each part of a fragmentary, intricate nebula too can be viewed as a separate nebula.

The Orion Nebula, the brightest one, is generally taken to be a glow, about one degree across, that envelopes the four O stars called the Orion Trapezium. But faint nebulous regions extend much farther and embrace almost the entire enormous constellation Orion.

Vast glow regions with uncertain boundaries are often to be found in the Milky Way and are called hydrogen fields or H II regions, since their glow is chiefly due to recombination of ionized hydrogen, like in planetary nebulae.

Hubble has shown long ago that the glow of gaseous nebulae is mainly caused by irradiation with ultraviolet radiation of O and B0-B1 hot stars, but no cooler. Since the temperature of these stars is lower than that of most of nuclei of planetary nebulae, ionization and excitation in them are lower: bright ultraviolet lines $\lambda\lambda 3727$-3729 of oxygen are



Rosette Nebula (NGC 2237) on 15/16 January 1972. In the centre there is a young hot star of high luminosity.

intense, but green lines of oxygen are faint.

A star (or a group of stars) that excites glow may lie either inside or at the edge of a nebula, or even outside it. For this reason, and sometimes also because it is far away, it is not possible to establish what star causes the nebula to glow. Such stars have not been found yet for a number of filamentary nebulae, whose glow may even be of some other origins.

Diffuse nebulae and hydrogen fields glow so faintly that their spectra can only be obtained using especially fast nebular spectrographs. Their radial velocities are of the same order as those of the stars that illuminate them, but it is possible that the mutual connection of a nebula with the star is temporal and random, and not of genetic nature, as is the case with planetary nebulae and their nuclei, which have high peculiar velocities reaching 200 kilometres per second.

Diffuse nebulae have lower velocities, so suggesting that they are involved in the revolution about the Galaxy's centre in the plane of the Milky Way in near circular orbits, whereas planetary nebulae seem to have more extended orbits and high chaotic velocities.

The aggregate of diffuse gaseous nebulae and hydrogen fields form a fragmentary layer of gas with a thickness of about 600 light-years in the plane of the galactic equator. The layer coincides with that of hot giants without which the gas clouds would not glow.

A hot giant inside a cloud makes it glow only in accordance with the size of the respective Strömgren zone (the zone of complete ionization of hydrogen). Outside it gas is invisible and most of light nebulae seem to be surrounded by zones of invisible neutral hydrogen. It is believed that they are arranged along spiral branches, as is the case with diffuse nebulae visible in the nearby spiral galaxies of later types and in the Galaxy. Astronomers therefore try to establish the localization of spiral branches of the Galaxy from the arrangement in it of hot giants and diffuse nebulae. But it is often overlooked that this evidence is not independent, since the distance to a nebula is taken to be the distance to the stars that excite their glow and sometimes perhaps erroneously assumed to be such. There is no other way of determining the distance to diffuse nebulae.

The distance to hot stars is estimated fairly crudely from comparison of the nominal absolute magnitude with the apparent magnitude. Absolute values are as yet established fairly approximately. We should also make allowances for the interstellar absorption of light near the galactic plane and over large distances. The data available are yet incorrect. Some difference in the spatial distribution of hot giants and diffuse nebulae consists in that there is sometimes no nebulae at giant clusters.

Masses of diffuse nebulae are worked out from the "emission measure", which is the product $n_e^2 \times l$, where n_e is the electron concentration, and l is the expected length of the nebula in parsecs. Accordingly, this quantity, which varies with the surface brightness, gives the number of hydrogen atoms along the ray of vision within a column with a cross-sectional area of 1 square centimetre and length equal to that of the nebula.

Having calculated the electron temperature or taken it to be 8,000 K we can obtain n_e from the emission measure substituting the assumed value of l. Glow can be found that have an emission measure

Orion Nebula (M 42) on 9/10 February 1975.

as low as several tens. Densities of diffuse nebulae usually appear to range from tens to hundreds of electrons (protons) per cubic centimetre, and at the centre of the Orion Nebula it is upwards of 1,000, but generally their densities are lower than those of planetary nebulae. In hydrogen fields the density drops to $n_e = 1$.

Multiplying the proton mass by n_e and by the nebula volume (sometimes assumed), we obtain the nebula mass. The first such determinations have been carried out by the author and his co-workers for the Orion Nebula and Omega Nebula giving 166 and 515 solar masses, respectively. It was later found that masses of individual nebulae range from 0.1 to hundreds of solar masses, and masses of complexes may be as high as thousands of solar masses. The masses of the smallest diffuse nebulae are close to that of planetary ones. The sizes of diffuse nebulae range from fractional parsecs to tens of parsecs.

In addition, gaseous nebulae sometimes show a continuous spectrum. Sometimes it is undoubtedly due to dust, especially when some dark streaks are seen against the background of a nebula, as is the case with the Trifid Nebula. The Orion Nebula contains much dust, which follows from the fact that the hot stars within it are strongly reddened. With such a dust density over a parsec it would absorb about 10 stellar magnitudes.

Some nebulae are more dusty, others less so, sometimes one part of a nebula is dusty and another is gaseous. The absence of traces of gaseous spectrum in many dust nebulae by no means suggests that they contain no gas. B1 and later type stars illuminating them cannot produce the required ionization and glow of the gas, which is still scarce in dusty nebulae, since calculations show that even at $n_e = 10\text{-}15$ the B2-B3 stars would produce a noticeable gas glow. But the opposite is not clear either: why are there no purely reflective nebulae illuminated by O and B0 stars?

According to Shain and Pikelner, the continuous spectrum in many gaseous nebulae is due not to dust but to two-quanta transitions, just like in planetary nebulae, whereas this spectrum has earlier been ascribed to dust. It may well be that there is some dust in bright gaseous nebulae, but it glows by reflected light so faintly that its continuous spectrum cannot be distinguished against the background of the bright spectrum caused by two-quanta transitions in gas.

The large masses of diffuse nebulae send out fairly noticeable thermal radiation.

Many research efforts are now devoted to gas-dynamics studies of the fate of diffuse nebulae. Large amounts of cool gas can of course be chained by gravity, but in the Galaxy everything is in motion.

Insufficient knowledge of density distributions and other conditions in real nebulae, their diversity and the differences in formulation and solution of theoretical problems, all these do not enable as yet unequivocal conclusions to be reached as to whether or not diffuse nebulae are being dispersed or they on the contrary are being condensed. Observational evidence available, too, gives no definite answer to this quiestion. According to certain authorities, cool gas can condense into stars and dust if any condensation nuclei are available in the form of complex heavy molecules, and so on. Hot, ionized gas condenses by no means.

Dust particles collide with one another and atoms of cool gas and can thus either merge and grow or evaporate, which also influences the density of the surrounding gas. We thus get an exceedingly complicated picture, which is dominated by the invasion of some dark matter into light regions of ionized gas. In the process the periphery of the dark mass begins to glow more intensely, forming a light, sharp fringe about its edge that is always facing the star. Especially narrow wedges of dark intrusions are called "elephant trunks".

The ionized gas density in the light fringe is much higher, and the dark region contains a mixture of cool gas and compacted dust. Theoretically, it is argued that when the hot star illuminates the cool gas ionization propagates faster in it than the pressure wave of gas being heated. The light fringe results when the ionization front comes to the dense cloud from the hot star. If the front encounters a high-density region, the latter remains unionized and the front circumvents this fluctuation. It is this that is responsible for the inclusion of H I regions into H II regions in the form of elephant trunks. The compression of cool gas in an elephant trunk by the pressure of the gas in H II region may result in a complete isolation of the gas cluster and give rise

to a globule. The compression of globules by hot gas and the formation in them of the so-called cumulative convergent shock wave make their gravitational condensation easier.

Filamentary nebulae of roundish configuration as a whole, such as Network in Cygnus, are a special case. But they are few in number and seem to be a result of supernova explosions. But filamentary nature is often inherent in the nebulae extended along the Milky Way. That they are extended cannot be explained by the action of the difference in the velocity of revolution of nebulae around the Galaxy's centre at different distances from the latter. The extension appears to be conditioned by the behaviour of the magnetic field of the Galaxy, whose lines of force lie in its plane and along the spiral branches.

Shain found some supporting evidence for this assumption by correlating the directions in which the nebulae extend with the data on the polarization of stellar light. Magnetic field allows gas to move freely along its lines of force but hinders the motion across them, and so it does expand along the lines of force, i.e. along a spiral branch. It seems that the constraining action of magnetic field, the concentration of lines of force at some places and their rarefaction at other places dictate the filamentary structure of large nebulae extended along the Milky Way. Ionized conducting gas retains field lines and entrains them. With intense chaotic motions the field lines become entangled together with gas fluxes, the field strength increases and the gas fluxes become more dense, which apparently results in the filamentary structure of vast expanses of gaseous nebulae, such as ones in Cygnus.

Planetary Nebulae

Thorough examination of good pictures of some planetary nebulae reveals that some of them appear as uniformly glowing or patchy disks, and others have the shape of a ring or a ring against the background of a disk. More complex or even mysterious forms are less frequent, but as a rule a planetary nebula is symmetrical and clearly delineated. The angular diameter of the largest planetary nebulae is about half of the Moon's, i.e. 1/4 degree. Some of them, the most remarkable ones, were given funny names because of their resemblance of something: Owl, Eskimo, Saturn. Most planetary nebulae are so small that they cannot be distinguished from a star even through the largest telescopes. How then are they found? Spectrally.

An overwhelming majority of stars have continuous spectra with dark lines. In any event, they contain conventional lines of the known chemical elements. But the spectra of gaseous nebulae are spectra of thin gases: they contain the so-called forbidden lines not to be observed in terrestrial laboratories that only occur in exceedingly tenuous gases under the condition that the gas is exposed to weak light fluxes. This has already been discussed earlier in the book.

The first to be observed among the forbidden lines only occurring in gaseous nebulae were the brightest: two green lines, which were assigned to an unknown gas that is only present in nebulae. Since it was called nebulium, its lines were referred to as nebular. The green nebular lines in planetary nebulae are brighter than the blue-green line of the Balmer series of the hydrogen H_β. This is their telltale sign.

Before World War II only about 150 planetary nebulae were discovered. Now we know more than 1,000 of them.

In 1887 the Irish astronomer Dreyer compiled a catalogue containing almost 8,000 stellar clusters and nebulae. Nebulae are often designated by numbers from this catalogue, for example, NGC 6720 where NGC stands for the New General Catalogue of Dreyer. A supplement to it published in 1894 and 1908 was designated by IC, i.e. Index Catalogue.

The Dreyer Catalogue mainly contains the galaxies that could not be identified at the time. It is utterly inconvenient to look up there the few planetary nebulae, besides it provides no necessary and already known data. Also, it does not contain a multitude of nebulae discovered later. This led the author to compile since 1931 three catalogues of planetary nebulae carrying all the important data on them: the position in the sky, size, brightness, physical properties, and so on. The last catalogue contains about 600 nebulae, at present however their number came to more than 1,000. Some planetary nebulae are designated in a different way.

Dumb-bell Nebula (NGC 6853).

How are planetary nebulae discovered? It is hopelessly tedious to photograph the spectrum of each and every faint star using a conventional spectrograph in order to find out whether or not a nebula is planetary. All the earlier found planetary nebulae have the total brightness of the stars from the 7th to the 14th magnitude. There are already about 15 million stars of the 15th magnitude, and fainter stars are even more numerous.

They often discover planetary nebulae using the objective prism. It is a prism with a refraction angle of 3-7° that is placed now in front of a fast telescope with an aperture of 25-60 centimetres. It embraces a patch of the sky about $3 \times 5°$ and gives the spectra of all the stars that at a given exposure leave their traces in the patch of the sky. Filament spectra are thus obtained of hundreds of stars, among which the spectra of planetary nebulae are immediately distinguished by their bead-like appearance. Beads are the monochromatic images of the planetary nebula at nebular, hydrogen and other lines. Faint planetary nebulae give only one image of the brightest line, which is generally either the principal green line of nebulium or the red hydrogen line H_α. A great many such nebulae have been discovered at the Abastumani Observatory in Georgia, USSR.

Large nebulae with low surface brightness are discovered from pictures obtained using fast telescopes with a red filter, which eliminates the dark sky background that interferes with the identification of nebulae.

The small scale of pictures of a majority of planetary nebulae only enables their position in the sky and at best their total brightness to be found, if only with an insufficient accuracy. But we know nothing about their diameter, structure, spectral lines, and all our knowledge is based on observations of about 50 of the largest and brightest objects.

At the centre of sufficiently large planetary nebulae a faint star is normally seen. As a rule, it is fainter than the total brightness of its nebulous envelope. Now that we know the cause of nebular glow we can argue that each of them has a star, a nucleus. It is only invisible because its brightness is weak. Nuclei are even more difficult to study than nebulae themselves, since it is only rarely that they are brighter than 10^m, more often 16^m-18^m or invisible at all.

Spectra of nuclei are of three kinds: O-type with dark lines, Wolf-Rayet with bright lines, and continuous without any lines. The latter may be caused either by extremely high temperatures or strong Stark effect (line broadening in interatomic electric fields), if the atmospheres of nuclei are very thin and highly compacted. Judging from the type of their spectrum, nuclei are as hot as the hottest among normal stars. Nuclei are so far insufficiently studied. According to half-hypothetical calculation carried out by the author, based on statistical distributions of planetary nebulae in space, the mass of a nucleus is on average about two solar masses. This is far less than the mass of a normal O-type or even Wolf-Rayet stars.

A remarkable phenomenon has been found in the giant nebula NGC 7293, photographed with the help of a 5-metre telescope in hydrogen red line. It is probably the nearest one, its distance being hardly more than 100 parsecs, or 300 light-years. It is for this reason that only in it hundreds of finest filaments have been found that are directed strictly radially to the nucleus. These filaments account for the inner, brighter part of its ring, but they are also observed inside it against the dark background of

the inner ring, at a large distance from the nucleus. Absolutely straight filaments are about 1″.5 thick, a limit of the telescope's resolving power, and about 1,000 AU long. These finest filaments are, however, enormous if we take into account the distance of the nebula. They are twice as thick as the diameter of Pluto's orbit, and about light-month long. The nature and origins of the filaments are absolutely unknown so far. Undoubtedly, they are directly involved in the formation of the envelope and are somehow associated with its nucleus.

In general, planetary nebulae have a simple configuration and a clear edge. But there are some exceptions. For instance, the nebula NGC 2440 is intricately chaotic. There are many tattered filaments at its edge. Lengthening the exposure turns it into a substantially larger and more regular nebula resembling a butterfly in appearance. But an overexposure turns it into an almost regular ellipse with an absolutely sharp and bright edge.

The nebula NGC 6720 in Lyra has been known for centuries as a ring nebula with a sharp edge. The pictures taken in 1964 show that it has a second, outer, extremely weak envelope with an uneven edge, and a third, still fainter envelope. As a result, the diameter of the nebula "became" 2.5 times larger.

In earlier times some faint appendages have been found in certain planetary nebulae, sometimes in the form of thin, faint straight or elliptic filaments, and sometimes in the form of spiral filaments, just as with NGC 650-1. At a short exposure it looks like an irregular quadrilateral, and at a long exposure the filaments at its edges look like "handles" attached to it. The appendages and filaments are less ionized than the main bulk of the nebula. They mostly radiate at hydrogen and ionized oxygen lines.

Worth especial mention is the distinct filamentary structure of a number of disk shaped nebulae. The filaments are short, they look like worms and correspond to the local compressions of gas. In the gaps between them radiation is weak. Through them the radiation of the nucleus can "leak" into space without being used for the nebular glow. This makes the determination of the true average density and the nebula's density more difficult with the use of certain methods. When at some part of the nebula its density is twice as high, then the radiation at forbidden lines is twice as high too, and at hydrogen lines four times higher. It is argued that the nebula's envelope is 30-70 per cent gas-filled, but in various nebulae this quantity must be different.

Expansion of Planetary Nebulae

One of the most important properties of planetary nebulae is their expansion, which has been discovered by spectral analysis. When a spectrograph has a sufficiently high dispersion and a spectral line is fairly wide, we can study the structure of a planetary nebula. In the process much additional evidence is revealed.

If the spectrograph slit completely covers the nebula's image, then the spectral line appears to be split in the middle. As the slit shifts farther away from the centre, the splitting is reduced, and at the disk's edges both line components merge. It follows from the small width of the split line that in the gas layer that forms the envelope the velocities of molecules correspond to thermal velocities and that there are no noticeable turbulent motions in the nebula. But the splitting of all spectral lines at the middle can only be accounted for by the radial expansion of the nebula that is essentially an envelope hollow inside and transparent to the self-radiation. This transparency is due to the extremely low density of the envelope. The transparency of planetary nebulae follows from the following example: distant galaxies are seen through the giant planetary nebula NGC 7293 (Helix) in Aquarius.

In expansion, the centre of the hemispherical surface facing us approaches us and its radial velocity equals that of expansion. At the same time the centre of the opposite hemispherical surface recedes and its velocity along the line of sight also equals the expansion velocity. At this point the spectral line is maximally shifted to the red end of the spectrum, whereas the line point is maximally shifted to the blue end of the spectrum from the centre of the front hemisphere. Accordingly, the half of the distance between the split line components, i.e. the half-width of the entire line, corresponds to the true expansion velocity. According to Doppler's principle, from this half-width of the split line represented in terms of the wavelength difference in angstroms, we can work out the velocity in terms of kilometres per second. Farther away from the visible

centre, the expansion velocity is directed at an angle to the line of sight. Its projection on the line of sight is smaller and the spectral line is less shifted from the standard position.

At the outer reaches of the visible disk of the nebula the expansion velocity is normal to the line of sight and appropriate points of the line are at their standard position. True, these points too are shifted by the amount corresponding to the velocity of motion along the line of sight of the entire nebula as a whole. These radial velocities are also important for studies of nebulae, all the more so that being so far away from us they almost do not move about the sky, and their total (spatial) velocity is determined inaccurately. If the nebula were not absolutely transparent to self-radiation because of its extreme thinness, we would not observe the red component of the split line formed by the receding nebular hemisphere.

When the expansion velocity is small and the amount of line splitting is only slightly larger than the component width governed by the thermal chaotic motion of atoms in the envelope, then instead of the distinct splitting we only observe line broadening. If the widths or densities of the front and rear hemisphere differ significantly, then the intensities of the split line components differ markedly.

Note that it has not been found for sure that the lines are generally inclined, so suggesting that planetary nebulae have no noticeable rotation about their axes. Should it be the case, then, according to angular momentum conservation, at the earlier stages of expansion when the nebula was thousands of times smaller than now, its rotation would be so fast that it could not exist at all.

As for the expansion velocities themselves, in the cases studied they were 15-30 kilometres per second, reaching a maximum of 53 kilometres per second with NGC 2392. It is impossible to find the expansion with nebulae of small angular size.

The general picture of expansion deduced from the form of spectrum might have more complicated details. In some nebulae, especially in IC 418, the expansion velocity generally decreases with ionization potential of given atoms, i.e. with energy required for their ionization. Whereas some gases quickly expand with velocities up to 23 kilometres per second, others, e.g. hydrogen, do not expand at all. This rule too has some exceptions. For example, some atoms with the same ionization potential as hydrogen recede from a star at 10 kilometres per second. With other planetary nebulae all the gases in the envelope recede at the same velocity. These variations from nebula to nebula and from one species of atoms to another, together with insufficient knowledge of their exact relative distribution in the envelope, make the definitive explanation of these facts impossible. The facts just described seem to be associated with the different light pressures for different atoms, with the degree in which they are mixed up or separated in space, with the temperature of the nucleus, and perhaps with the initial conditions of their expansion.

It is only natural to desire to check directly the expansion of planetary nebulae that has been found from spectral analysis. This would require finding the increase in the angular diameter of nebulae. The best results have been obtained by Liller and his American team. They obtained pictures of some nebulae where according to the calculation of the author made in 1948, the most noticeable expansion could be expected. They compared their pictures with those taken using the same telescope some 60 years back.

In eight cases the angular expansion appeared to be less than it had been expected and in six cases it was absent at all and could not be accounted for by an error in the estimation of the distance of the nebulae but could be explained by the assumption that the nebula density remained the same due to ejections of matter by the nucleus (note that errors, unavoidable in measuring any small quantities, would rather give an overestimation of changes in nebular diameters, not an underestimation). In one case, it seems, the observers found agreement between angular and linear expansions and a variation of expansion with distance from the nucleus. According to another estimate of the distances, again the hypothesis is required that the envelope's mass is made up of continuous effluents of gas from the nucleus.

Although the masses of nebular nuclei are unknown, and their distances, and hence envelope sizes, are only known inaccurately, the following is beyond doubt. When expansion occurs at 10

kilometres per second and more, this velocity is higher than that of separation of the envelope from the nucleus. The attraction of the nucleus cannot prevent the expansion and it proceeds essentially with a constant velocity. It is also beyond doubt that with constant velocity already in several tens, at most hundreds, thousand years the envelope of a planetary nebula will expand so that it will become invisible and scatter in space. Thus, as the author noted back in 1931, planetary nebulae are suppliers of diffuse gas into space.

Another conclusion is also of interest. It is quite obvious that 10^4-10^5 years ago, when it just separated from the nucleus, an expanding envelope was the size of a star. Consequently, from the astronomical point of view planetary nebulae are very young formations, they have come into being fairly recently. They are younger than supergiants, which are estimated to be 10^6-10^8 years old. Additionally, planetary nebulae are ephemeral, they are short-lived. These cosmic moths live no more than 10^5-10^6 years. This does not mean, however, that their nuclei are ephemeral too. These may be old stars, which will glow, without changing noticeably, for a long time after their envelopes have dispersed. Sometimes, however, interstellar medium impedes the expansion of a nebula or the nucleus makes up the gas in the envelope, then the planetary nebula may live longer.

Evolution of Planetary Nebulae and Their Nuclei

The temperature of nuclei of planetary nebulae cannot be determined by techniques used with normal stars, because their spectral lines are often either too bright and wide or too faint or not seen at all. Also, the temperature of hot nuclei cannot be found from energy distribution in the continuous spectrum, since in the visible range the distribution varies only slightly with temperature.

Zanstra indicated a possible way of deriving the temperature of a nucleus. His idea is that the brightness of a nebula at visible lines corresponds to the energy in the far ultraviolet spectrum of the nucleus, where quanta are sufficiently energetic to ionize the atoms of the nebula's envelope and tear off electrons in collisions. So from the brightness of the nebula at visible lines of hydrogen it is possible to determine the brightness of the far ultraviolet range of the nucleus's spectrum with a wavelength shorter than 912 angstroms (quanta with longer wavelength are no more able to ionize hydrogen). Comparison of the number of these quanta with that in the visible range of the nucleus's spectrum enables now its temperature to be determined more exactly, if the nucleus is a black-body emitter (in which case the energy distribution over the entire spectrum depending on temperature is known theoretically).

Recently Khromov derived the energies of three points of the ultraviolet spectrum of a nucleus from the sizes of ionization zones of various atoms. From these energies and using Planck's formula, he estimated the temperature that characterizes the ultraviolet range of the spectrum to be about 150,000 K. In the more long-wave region the nucleus's spectrum is represented by Planck's formula for lower temperatures. In 1965 in the author's laboratory a good agreement was found between the visible part of the spectrum of a dozen nuclei with the Planck formula at temperatures from 15,000 to 65,000 K.

The issue of the temperature of nuclei is as yet poorly studied. Obviously, it has a high dispersion, because nuclei with absorption spectra O9-O5 have the same temperatures as normal stars of the same type, i.e. not higher than 35,000 K.

Many years ago the author found that nuclei with the Wolf-Rayet spectrum have higher temperatures than those with O9-O5 type spectra. The highest temperatures are found with nuclei with a continuous spectrum having neither dark nor bright lines. The former agrees with what we have for normal O and Wolf-Rayet stars, and the hot stars with a continuous spectrum, apart from nuclei of planetary nebulae, are unknown, save for two or three white dwarfs.

Whatever the final data on the distances of individual planetary nebulae, the conclusion still holds that was made 40 years ago, when the distances of these objects were first estimated. It states that luminosities of nuclei are on average much lower than of normal hot stars with the same spectra and temperatures but devoid of extended nebular envelopes. Moreover, considering that the luminosity of nuclei has undoubtedly high dis-

persion, it can be said that at least some of the nuclei are similar to white dwarfs of Sirius's companion type.

Nuclei are yet more similar to former new stars in temperature spectrum and luminosity. It would even be more appropriate to call them, as was initially supposed, blue or ultraviolet dwarfs. Their masses cannot be less than the Sun's, and luminosities of some of them are much lower than the Sun's, therefore at higher temperatures their volumes are far smaller than the Sun's, and the densities are enormous, approaching perhaps those of white dwarfs, although maybe not reaching them.

We are not going here to provide numerical data on their physical characteristics because of their insufficient reliability. Specifically, we do not know the corrections to be added to their visual luminosity in order to obtain their bolometric luminosity expressing their total radiation at all wavelengths. This is because their radiation does not obey the laws governing black-body radiation. Bolometric luminosities of nuclei, it seems, have much lower dispersions than photographic luminosities.

With high gravity stresses on the surface of white dwarfs their spectra are known to show line shift to the red end as predicted by Relativity. Such an effect can also be expected with dwarf nuclei of nebulae. To find it, it is necessary to compare the wavelengths of the nucleus's spectrum with the wavelength of the ends of the same spectra of envelopes (since at the middle the lines are split by broadening). Such a comparison is difficult to effect practically. In two cases the result appeared to be negative and in one nebula the red shift in the nucleus has been measured, but the data obtained do not merit high credit. A substantial red shift is not obligatory, since the lines of the nucleus may also be produced at a large height in its atmosphere where the gravity stress is less and the nucleus may have a not very small radius.

In two cases the author has found changes in spectral line intensity, which can be accounted for by nucleus temperature changes alone that seem to have temporal nature. This also suggests the possibility of fast evolutional variations in nuclei. Such variations, if any, exert a substantial influence on the evolution picture and nuclear envelopes, the questions to be discussed later in the book. Neither brightness fluctuations nor duality of nuclei have as yet been found.

Looking for the causes of expansion and possible time variations constitutes a complex problem. It has been assumed that it is the light pressure from the nucleus that is responsible for the expansion (it is different for different nuclei). Now these arguments seem to be doubtful.

The role of the gaseous pressure seems to be more significant. The expansion velocities are close to thermal ones and to the gas expansion velocities in a vacuum. Gurzadyan was the first to introduce the hypothesis that the magnetic field inside the nebula markedly affects the distribution and motion of gases. His theory met with many criticisms but we believe that without the assumption of the magnetic field it is impossible to explain many features of the structure of planetary nebulae.

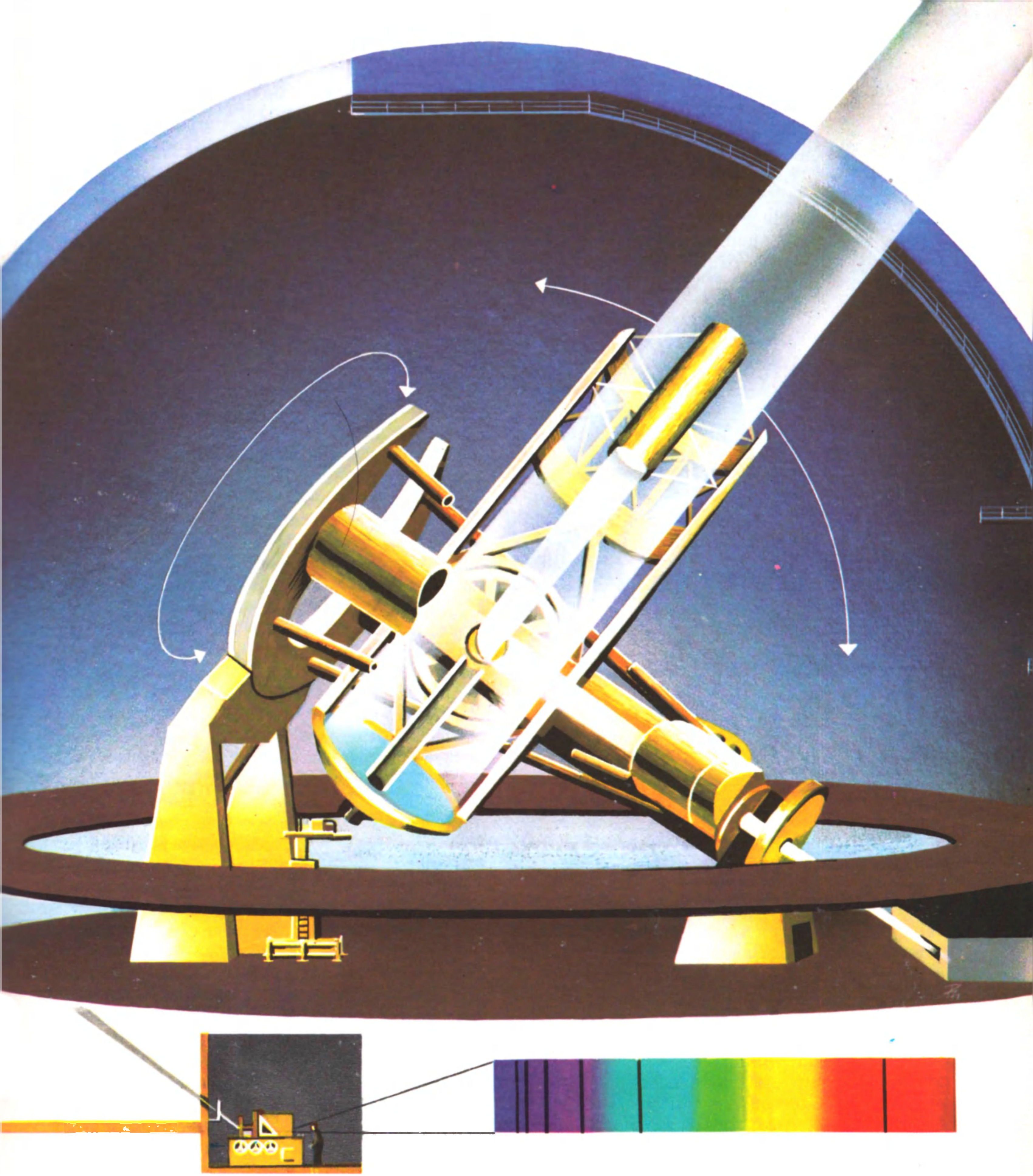
It is normally assumed that a planetary nebula expands indefinitely, in the process its luminosity and surface brightness quickly falling off. In several tens, or hundreds, thousand years it is seen no more and its envelope scatters in the surrounding interstellar space.

Only the nucleus remains – a blue dwarf, unless by the time it changes its behaviour. We have already noted above, however, that the drag of the interstellar space and other factors can slow down the evolution rate, but there is no saying to what degree.

Significantly, the expansion of a large number of planetary nebulae is a directly observable act of scattering of gases that earlier belonged to the nucleus, i.e. to stars.

Up till now it is absolutely unknown which stars and why produce planetary nebulae. Any attempts to see in a pre-nebular stage some types of cool variables, nova explosions or gas ejection by Wolf-Rayet stars are as yet a failure. In the last two cases the velocities of envelope ejection or constant gas outflow are far too high to produce a slowly expanding planetary nebula. But it is beyond doubt that during the Galaxy's lifetime the planetary nebulae have filled it with an amount of gas that

The cross-section through the 5-metre mirror telescope at the Palomar Observatory, USA, used above all for spectroscopic studies.



accounts for a noticeable proportion of all gases observed in the Galaxy now. And until now there are two conflicting hypotheses: that stars develop from clusters of diffuse matter and that on the contrary, they develop from some superdense matter. By the way, that diffuse matter is produced by stars, if only in part, is now a commonly observable fact.

Interstellar Gas

Gas, everywhere gas! Collected in gigantic spheres, it forms countless stars which contain the greater part of matter in the Universe. Tenuous cool gas, filling enormous spaces in the form of gaseous nebulae, constituting the envelopes of numerous stars and the atmospheres of the planets. And all of this in airless space. But is it really airless?

Our concepts of a vacuum and airless space are relative. In an old-type electric bulb there "was no air"; we therefore say that it is evacuated. It contained a vacuum in comparison with the air in the room. But the physicist, by the use of his best pumps, can evacuate air from any glass tube so that by comparison it would seem that the space within the bulb is swarming with myriads of molecules.

Gaseous diffuse nebulae, with their density of the order of 10^{-19} grammes per cubic centimetre, are scattered in airless space. But as we shall see, this space is not entirely a void; it also contains gas, and between any two stars there is a gaseous medium, however tenuous it may be.

But what is this gas? To be sure, this is not terrestrial air, even rarefied. The history of study of this gas has brought us much interesting and unexpected.

In 1904, while studying the spectroscopic binary δ Ori, to obtain greater accuracy of determination of its radial velocity, Hartmann measured the position in the spectrum of all the dark lines visible in it. Indeed, if a star moves as a whole in its orbit around the centre of mass of the system, *all* the lines of its spectrum should shift by the same amount so that within the measurement error the shift of any line should correspond to the same velocity of approach towards, or withdrawal from, us. We already know that with such a *periodic* orbital motion the spectral lines *periodically* change their shift. At the centre of δ Ori all the lines behaved properly except the lines of ionized calcium. These two lines for some reason did not participate in the general periodic oscillations of the position of spectral lines, but stubbornly stayed put. Whether the star was approaching or receding made no difference to the calcium lines.

The stubborn lines belonged to calcium atoms, and so Hartmann could draw no other conclusion than that for some reason or other calcium does not participate in the orbital motion of the star. Since the calcium lines are visible as dark lines (in absorption), it is obvious that the light of the star passes through it and is absorbed in it, but this element is not found in the atmosphere of the star, which causes the other absorption lines. The atmosphere moves together with the star, but the calcium does not move with it. Our binary may be submerged in an extensive cloud of rarefied calcium, in which it moves without entraining it.

This type of calcium line has been called stationary, i.e. unchanging or fixed. Later on, the stationary lines of calcium were found in the spectra of many other spectroscopic binaries, but only when the stars were of the early spectral class B.

Slipher, however, found it more probable that the stationary lines are produced not by a cloud of calcium in which the star is submerged, but by clouds of calcium or its continuous mass situated along the entire line of sight between the star and the Earth. In other words, the calcium is not in the neighbourhood of the star, but is an interstellar gas. This conclusion has been confirmed. The term "stationary lines" was then replaced by the term "interstellar lines".

This was found in the following way. When it became known that the temperature of a stellar atmosphere determines the form of its spectrum, it became possible to determine the intensities of the different lines produced by the atmosphere of a star of a given chemical composition and temperature. It was found that such hot stars as B-type stars do not have in their atmospheres any atoms of ionized calcium, since the atmospheres are too hot to contain them. All the calcium there is doubly ionized and the spectrum cannot have its lines. This means that the ionized calcium responsible for the stationary lines should lie far away from the star where it is no longer so hot and calcium can exist.

It was then found that it is not only spectroscopic binaries that reveal these lines; they are present in the spectra of most hot individual stars. The lines cannot in general be called stationary, since a single star does not experience orbital motion. Such a star moves relative to us constantly at the same velocity and therefore all its spectral lines are displaced equally. It was found, however, that in such hot stars the displacement of the lines of ionized calcium corresponds to a completely different velocity from the velocity with which the star itself moves.

If ionized calcium fills all of interstellar space, then its lines, which we will see always displaced in a particular way, should be present in the spectra of stars of any type. Unfortunately, the cooler stars themselves contain ionized calcium in their atmospheres, and hence its lines in their spectra. These lines are broad and strong and mask the thin and faint lines of interstellar calcium. In certain cases, however, it is possible to detect these thin "interstellar" lines, superposed on the broader "stellar" spectral lines.

A decisive development was a comparison of the intensity of the lines of interstellar calcium with distance to stars, made in Canada by Plaskett and Pearce. It turned out that the more distant the star, the stronger are its lines of interstellar calcium. But this was to be expected if calcium fills all the interstellar medium. The more distant a star is from the Earth, the longer is the path its light travels before it reaches the Earth and the more absorbing atoms of calcium it meets along its path. The greater the quantity of calcium atoms that absorb the light of a star, the more the light is attenuated and the darker and stronger will be the absorption line in the spectrum. We must agree with this explanation.

Moreover, we now have the possibility, by determining from observations the relationship between the intensity of the lines of ionized calcium and the known distances to stars, to work out from the intensity the distances to those hot stars for which. they are still unknown. Thanks to the interstellar calcium! This is what we should say in many cases, since frequently we have no other method of determining the distance to a particular star.

Plaskett and Pearce were also able to demonstrate that the interstellar calcium participates in that general rotation which involves all the stars of our stellar system. By comparing the radial velocities of stars caused by this rotation with the radial velocity of the interstellar calcium (on the basis of the line shift in the spectra of the same stars), we can see that the latter is only half as great as the radial velocity that should be obtained for a particular star in accordance with the theory of galactic rotation. But this lesser velocity relative to the Sun in the course of rotation of the Galaxy should be characteristic of a point one half as distant. One conclusion can be drawn: interstellar calcium participates in the rotation of the entire stellar system, together with the stars and in conformity with the same laws, since the centre of mass of the columns of gas that is situated between any star and the Earth in all cases coincides with its midpoint. This means that in the space between the stars the distribution of calcium is rather uniform.

But it was discovered later that, much like cosmic absorbing dust, the calcium is concentrated in the plane of the Milky Way. It was also found that it is not distributed as a continuous medium but rather in the form of numerous clouds. The dimensions of certain calcium clouds are as great as 2,000 light-years.

Until the properties of atoms had been studied thoroughly by physicists, the exclusive or at least predominant presence of calcium between the stars was the cause of perplexity. It then became clear that ionized calcium absorbs light primarily in those two lines that are situated in the easily observed part of the spectrum. The atoms of other elements absorb light either in very many lines, like iron, or in a spectral region (ultraviolet) that is inaccessible for study due to its total absorption in our atmosphere. Therefore, the lines of other interstellar atoms, if any, either cannot be discovered or they are less noticeable because their total absorption is broken down into many different absorptions, a little in each line. There is therefore no good reason to assume that ionized calcium is the only or the predominant gas in the interstellar medium – it only makes its presence known more conspicuously than the others.

It is nevertheless possible to attempt to find other interstellar gases, at least traces of them, for he who seeks always finds. In fact, special investigations of stellar spectra have found interstellar sodium, and in

recent years astrophysicists have also discovered neutral calcium, ionized titanium, neutral potassium and even iron. In addition, late in the 1930s there were also discoveries of interstellar molecules of neutral and ionized hydrocarbons CH and CH^+, cyanogen CN, NaH and also certain lines of still unknown origin. The total density of the absorbing interstellar gas is several thousand times less than the density of the light-emitting gaseous nebulae.

Atoms of gas that somehow got into the interstellar space are ionized and excited by light quanta radiated by stars. They occasionally collide with these quanta. We have said occasionally because only a very small number of these quanta at a great distance from stars pass through a square centimetre of surface. There are likewise infrequent encounters of ions with free electrons, more infrequent than in gaseous nebulae with their greater density. While an atom of ionized calcium is wandering in space, patiently awaiting an encounter with some wandering electron, it can collide with some stellar light quantum with a wavelength of 3,933 angstroms, which will excite it to a higher energy state. Being unable to withstand such an excitation longer than one ten-millionth of a second, the atom returns to its initial, normal or unexcited state. At the same time it radiates back the absorbed quantum of energy with a wavelength of 3,933 angstroms. But it sends the quantum out not in the direction from which it was received, but in some other direction. Thus, a calcium ion situated between the Earth and a star, trapping stellar light quanta travelling in our direction, will "toss them to and fro", scattering the light, and less light will reach us than would be the case if this obstacle was not in the way. As a result, the starlight at this wavelength is attenuated and we see a dark line in its spectrum. Other interstellar atoms behave in the same way.

Knowing the structure of atoms and their absorption capacity it is possible from line intensity to estimate their quantity along the line of sight to the star, and knowing the distance to the star, compute the density of the interstellar gas.

The first steps taken in this direction give a density of $4 \cdot 10^{-32}$ grammes per cubic centimetre for interstellar ionized calcium. The total density of the interstellar gas is considerably higher and according to Eddington's estimate is not less than 10^{-24} grammes per cubic centimetre. If this gas consisted of hydrogen alone, at such a density each cubic centimetre would contain one atom only, whereas in the same volume of air there would be 10,000,000,000,000,000,000.

In actuality, this is about how matters stand, since hydrogen is indeed the principal component of the interstellar gas. The next most important constituents are oxygen and sodium, but more than 90 per cent of the atoms of the entire interstellar medium are hydrogen, and this refers to cosmic dust and meteoroids as well. The latter appear to account for an insignificant fraction of the total mass of the interstellar medium, in which the lightest of the gases is the most important constituent.

Unfortunately, interstellar hydrogen in absorption cannot be detected by optical methods and it is unlikely that it will ever be detected because in most corners of our Universe the overwhelming number of hydrogen atoms is in an unexcited state and therefore absorbs energy in the invisible far ultraviolet region of the spectrum.

There is a certain hope of seeing the familiar lines of hydrogen in emission, not in absorption. They can appear when free electrons are trapped by hydrogen nuclei and return to the orbit closest to the nucleus, to that with the minimum energy, step by step, remaining for some time in the second orbit from the nucleus. Such cases will not be frequent, and the emission of the bright lines of interstellar hydrogen should be faint.

Using long exposures, Struve has succeeded in discovering the faint hydrogen emission lines in several extensive regions of the Milky Way. This is a signal in visible light from interstellar hydrogen, but the author thinks that in many cases we are dealing with the overlapping of large, distant and very tenuous diffuse gaseous nebulae. Being faint and indistinguishable separately, they appear as an indefinitely extensive hydrogen-emitting H II.

This is confirmed by the fact that in addition to the hydrogen lines, in these same regions of the sky it has been possible to detect the bright lines of forbidden nitrogen and oxygen, i.e. the ordinary spectrum of gaseous nebulae has been found. Furthermore, in these regions astronomers have found O-type hot stars which always excite the luminescence of gaseous nebulae.

However, not only the existence but also the spatial distribution and the velocity of motion of interstellar

hydrogen have now been established reliably from the radio emission of interstellar hydrogen.

According to the estimates made by Dunham and Struve, the density of the individual gases in interstellar space, determined from the intensity of both their absorption and emission lines, is as follows:

Hydrogen	$2.7 \cdot 10^{-24}$ g/cm^3	Calcium	$7 \cdot 10^{-28}$ g/cm^3
Oxygen	$2.3 \cdot 10^{-26}$ g/cm^3	Titanium	$8 \cdot 10^{-30}$ g/cm^3
Sodium	$4 \cdot 10^{-27}$ g/cm^3	CH	$2 \cdot 10^{-29}$ g/cm^3
Potassium	$7 \cdot 10^{-28}$ g/cm^3	CN	$1.5 \cdot 10^{-29}$ g/cm^3

On the basis of an analysis of the apparent motion of stars, it is impossible to assume a density of interstellar matter higher than $6 \cdot 10^{-24}$ grammes per cubic centimetre, and most likely it is a value coinciding with the above estimate. It is interesting that according to certain estimates the mean density of interplanetary space in the solar system, bearing in mind its meteoric matter, is $5 \cdot 10^{-25}$ grammes per cubic centimetre. This is even less than the density of interstellar space. According to Greenstein's estimate, the density of interstellar dust (excluding gas) is $2 \times \times 10^{-25}$ grammes per cubic centimetre. Therefore, with respect to its mass the interstellar dust is inferior to interstellar gases.

In 1932 the American radiophysicist Jansky found the radio emission of the Milky Way, which appeared to be remarkably strong in the metre range. It turned out that it comes from two sources, one of which is the cluster of many gaseous nebulae in the Milky Way. Only the closest and brightest of them are visible. The more distant ones are not seen because of light absorption by cosmic dust. But the dust almost does not stop radio waves, and so the radio emission from distant nebulae merges to produce constant radio noise along the Milky Way. Sky maps have been charted that show its "brightness" over the entire radio range.

The other source of radio emission is the bremsstrahlung of relativistic electrons in interstellar magnetic fields. In the mid-1960s it was proven conclusively that interstellar magnetic fields do exist. Cosmic rays too contain relativistic electrons. We have already mentioned that as relativistic electrons are decelerated in a magnetic field the bremsstrahlung is produced, specifically in the radio range.

Hydrogen is ionized by hot stars that are few and that form a relatively thin layer not filling it wholly. Farther away from this layer and in this layer but closer to the centre of our stellar system there are likewise no hot stars and ionized hydrogen. Hydrogen may occur there, but it will be nonionized. Shklovsky has predicted that neutral hydrogen must emit the 21-centimetre line and that it must be sufficiently bright to be detected by radio telescopes. This was soon confirmed by observations. So cool invisible neutral hydrogen became accessible for observation almost in the entire volume of the Galaxy. After all, the 21-centimetre line is not absorbed by interstellar dust.

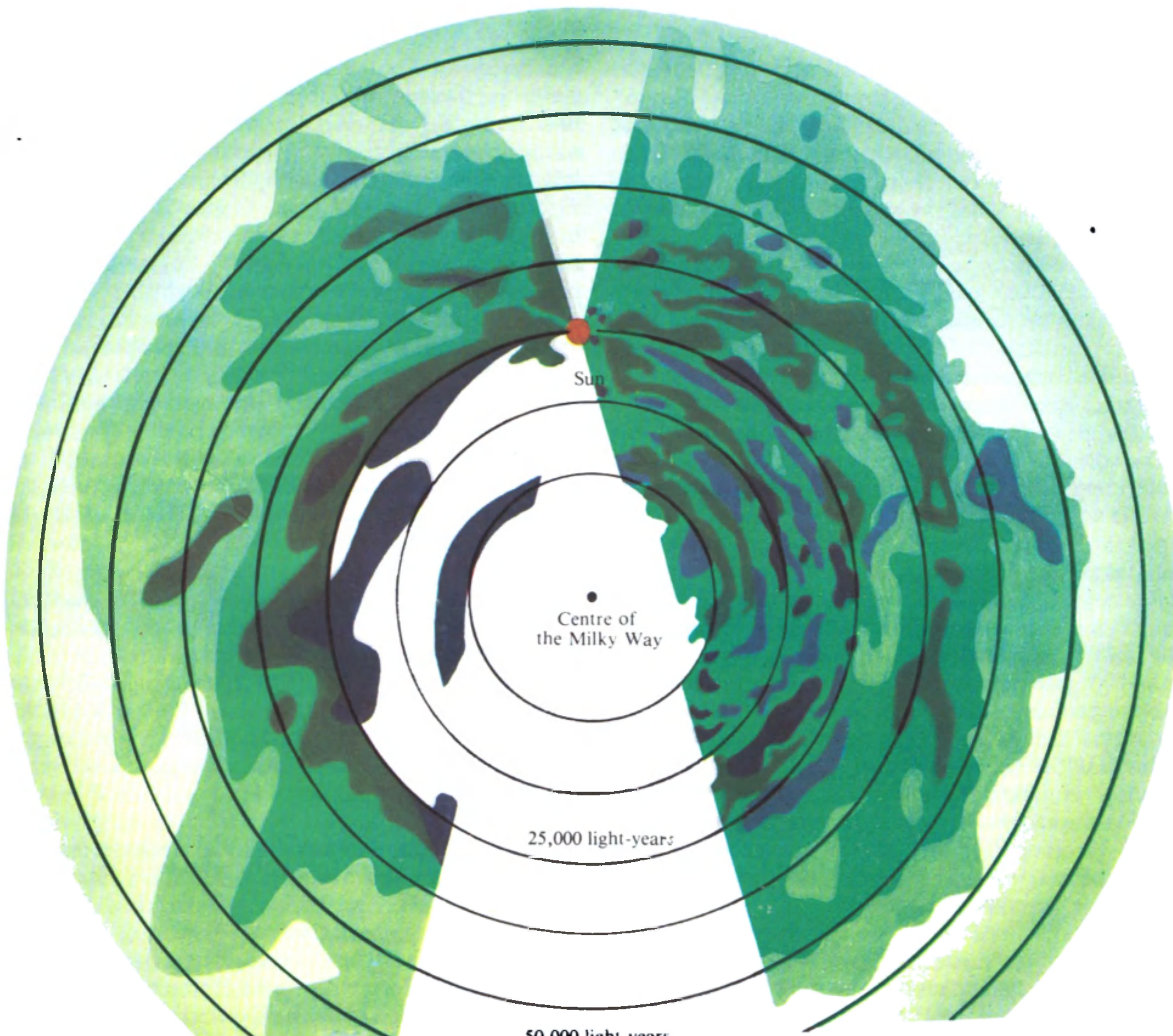
From the shift of a line emitted by a cloud of neutral hydrogen, we can establish the rate of the motion of the cloud along the line of sight. Knowing the law of rotation of the Galaxy and the velocity of the cloud, we can calculate its distance. From the line intensity we can find the density of clouds, and a knowledge of their distribution in space adds enormously to our understanding of the structure of our stellar system. In terms of mass, there is about as much molecular hydrogen in the Galaxy, as there is atomic hydrogen: about 10^9 solar masses.

Dust, interstellar gas, and hot galactic nebulae concentrate in a flat layer about 600 light-years thick, which is small in comparison with the size of the Galaxy. But individual clouds of hot and cold gases are also to be found at large distances from this layer where they move about chaotically.

In 1963 a radio telescope found an earlier predicted radio-frequency line of the hydroxyl OH in the interstellar space. The line is complex and includes several components, its wavelength being about 18 centimetres. It has been observed both in absorption and in emission, generally in the region of hot gaseous nebulae, but not all of them. The complex of OH-lines exhibited a number of as yet extremely mysterious phenomena. Specifically, the brightness was found to vary from day to day in a different way with different line components. Obviously, future development of science will explain these enigmas.

Infrared studies revealed interstellar helium, and in 1965-1966 it was also found by radio techniques. One of its major lines has a wavelength of 6 centimetres, and another lies near the hydrogen radio-frequency line with a wavelength of 21 centimetres.

By 1983, apart from atoms, all in all about 50 biatomic and multiatomic molecules were found in the interstellar space, mostly by radioastronomical



The distribution of interstellar hydrogen in the galactic plane, the average density being 0.8 hydrogen atoms per cubic centimetre.

methods. Among them, there are complex ones (containing up to 11 atoms) including water, ammonia, formic acid, and methyl alcohol. Also found were CO and CN typical of comets.

By 1980 it became possible to conclude that the percentages of various chemical elements in interstellar gas differ markedly from those in the solar and stellar atmospheres, although these atmospheres partly scatter gradually in space, and interstellar medium partly accumulates on stars being caught there (*accretion* by gas). For example, in some directions a shortage of many atoms has been found: in relation to hydrogen they are one third or less than in the Sun's atmosphere. Such anomalies are, however, of local nature.

How could it be that interstellar space had been filled with gas? Which is older – the diffuse interstellar gas and nebulae or stars? Later in the book we shall discuss these questions.

5. Islands of the Universe

Milestones and the Structure of Our Galaxy

When you gaze in the skies on a clear autumn night you notice the faint band of the Milky Way known to man since time immemorial. It owes its name to the fact that it resembles a faint dried band left by the spilt milk. American Indians believed that along it, as along a road, the souls of the dead rise into the heaven.

We are interested not only in the stellar inhabitants of the house we live in, but also in the architecture and dimensions of this house. We are interested in how the inhabitants are distributed and what areas are thickly and thinly populated by stars. Now looking into the starry wastes, into the star-studded heavens, we should endeavour to examine all of these. From the naive ancient picture of the world, assuming as reality the fact that seemed apparent to the eye that all the stars were at an equal distance away and all situated on the surface of a crystal sphere, we must move on to an understanding of the true spatial structure of the grandiose stellar system.

Scientific studies of the Milky Way and speculations about its structure contributed significantly to our understanding of the Universe.

The first thing that we seek to establish is the general outlines of our stellar system, if only in the roughest terms. Early astronomers succeeded in it long before the distance to the nearest star had been found. Originally, it was quite correctly assumed that the luminosity of all the stars is identical and the difference in their apparent brightness is exclusively dependent on their distance from the Earth. We now know that in actuality the luminosity of stars differs tremendously, but we also know that very bright stars are very few and that of the faint stars we can only see those that are close to us. Therefore, the majority of the stars that we can see will be average stars and the supposition is completely applicable to them *on average*.

Suppose that you were standing on a high hill overlooking a plain over which roves of old and young trees were scattered. They are different in height and you do not know the height of each. But by looking at them from the hill you can judge correctly about the distance to each group of trees from their *apparent magnitude*. Supposing you are the explorer of the Universe and the hill is our solar system, the trees are the stars. Apply this approach to the stars and make a visual evaluation. This method of study of the stellar Universe was proposed by William Herschel. By this method we can observe the positions of the stars in the sky and study the surface of the Moon and planets and also to scrutinize the motions of the members of the solar system.

A few words should be said about the life of Herschel. A musician, first serving in the Hannoverian army, he later moved to England where he became a musician to the king, but his evenings he spent observing the skies. After his discovery of Uranus he became very famous. Nevertheless, he still did not have the money to purchase a large telescope. He therefore began to make one himself. He was so good at this that he later constructed a reflecting telescope 120 centimetres in diametre, which for a long time was the world's largest. He made many discoveries with this instrument.

In order to clarify the outlines of the Universe Herschel began to count the number of stars of different brightnesses visible in the field of view of his telescope in different sectors of the sky – in the Milky Way and away from it. He discovered that the fainter the luminosity of stars, the greater is their number with approach to the Milky Way. The Milky Way itself, as Galileo discovered, consists of a myriad of faint stars, merging into a continuous shining mass and girdling the entire sky like a ring.

From these counts it became clear to Herschel that our stellar system extends farthest in a direction of the Milky Way in a plane passing through its mean line. Since the Milky Way girdles the entire sky, dividing it almost in half, our solar system is obviously situated near this plane (near the galactic plane, as it is called).

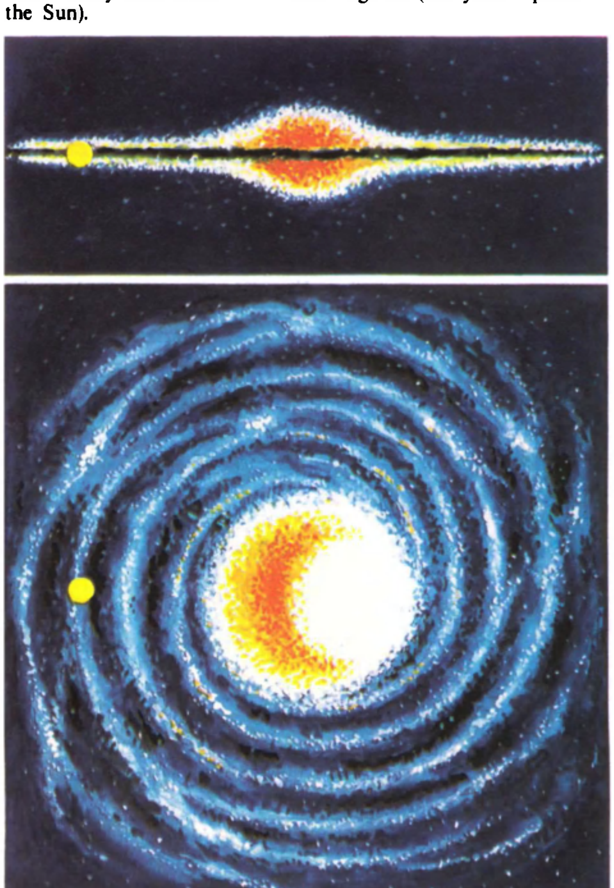
Herschel assumed, however, that with his giant telescope he penetrated to the limits of our stellar system, consisting of stars apparently uniformly distributed in space.

Struve, the founder of the Pulkovo Observatory, reconsidered Herschel's counts in 1847 and after studying the distribution of stars demonstrated the fallacy of such conclusions. Struve established that stars are distributed nonuniformly in space and are concentrated toward the plane of the Milky Way. He also demonstrated that our Sun by no means occupies a central position in this stellar system and that Herschel's very large telescope by no means reached its boundaries and it was therefore premature to try and

define the configuration of the stellar system. Herschel assumed that he, so to speak, was sitting with his telescope in the centre of a properly located grove, from which all outlying areas could be viewed. Struve demonstrated that Herschel was sitting somewhere in an immense forest, full of thickets and thinly wooded areas, and that the outskirts of the forest could by no means be seen.

The farther from the plane of the Milky Way, the fewer faint stars are visible and in these directions the smaller is the distance to which the stellar system extends. In general, our stellar system occupies a space resembling a lens or a lentil. It is flattened, being thickest in the middle and thinner toward the edges. If we could view it from above or below it would, roughly speaking, appear as a circle (not a ring!). However, from the side it would look like a spindle. But what are the dimensions of this "spindle"? Are the stars in it distributed uniformly?

The Galaxy seen from above and edge-on (the yellow point is the Sun).



This has been clarified in recent years, although the second question is answered by a simple inspection of the Milky Way, which has the appearance of a mass of stellar clouds. Some clouds are brighter and contain more stars (such as the constellations Sagittarius and Cygnus), whereas others have a lesser number of stars.

The apparent fragmentary appearance of the Milky Way is also due to the uneven distribution of clouds of cosmic dust, dark nebulae of different density that absorb the light of stars lying behind them. But even considering this the stellar Universe is nonhomogeneous. The Galaxy consists of stellar clouds and the solar system is situated in one of them, called the *Local System.* The greatest of the stellar clouds are in the direction of Sagittarius: this includes the Milky Way which is the brightest. It is the least bright in the opposite part of the sky.

From this it is easy to draw the conclusion that the solar system lies not in the centre of the Galaxy, which is visible from Earth in the direction of Sagittarius. Hence the Milky Way is a picture seen by us from inside the Galaxy, near its plane, but far from its centre.

In order to obtain a more correct picture we should now take into account the luminosity distribution of stars and the light absorption in space, which, as we have seen, is considerable and differs significantly in different directions.

To study the variation of stellar density with distance from us in various directions we count the stars on photographs of different sectors of the sky. In this process it is also necessary to take into account the fact that on each photograph a circular area in the sky corresponds to a volume *in space* enclosed within a cone with its apex in the solar system or, if you wish, in the observer's eye.

The study of the Universe is thus a rather complicated and tedious thing. To take account of light absorption it is necessary to determine spectral types and colours of a multitude of stars in each picture. Normal colours of stars in each of spectral types are known from the study of the nearby stars, whose colour is not influenced by light absorption. But the light of distant stars becomes the redder, the far-

Stellar clouds in M 17 on 9/10 June 1969.

ther from us and closer to the plane of the Milky Way they lie. The apparent brightness becomes weaker in proportion to its reddening. It is from the degree of reddening that the attenuation of the apparent brightness of each star is worked out. The necessity to photograph stellar spectra is an additional constraint as only the brightest stars and only distributions up to several hundred light-years can be studied. So we only study the closest neighbourhood of the Sun and only the inner part of the stellar cloud, of the Local System inside which we live.

To study, so to speak, the skeleton of the entire Galaxy, to define its shape, size and structure we make use of other techniques. The task of charting a region of steppe with sparse trees would be extremely easy, if throughout the entire territory we would encounter giant trees with inscriptions of the distances at which they lie. In the Universe, too, we have found giant stars which have a well-defined and well-known luminosity. Because of their great luminosity we can see them at an enormous distance. Among these the leading role is played by the variables, the Cepheids, which we can call the lighthouses of the Universe. Their luminosity increases, as we know, with increasing period of their variation. If we can determine the variation period of a star, we can say immediately what its absolute magnitude and luminosity are.

By means of observations it is easy to determine the variation period of a Cepheid and its apparent brightness. A comparison of the apparent brightness with the true brightness, i.e. with the luminosity L, immediately gives the distance to a particular Cepheid, since in transparent space apparent brightness varies inversely proportional to the square of distance. When light absorption in space is taken into account the matter becomes somewhat more complex. If we use the absolute and apparent stellar magnitude of a star, the distance, as we know, can be computed using a simple formula. From this formula we can find the logarithm of distance in light-years:

$$\log D = \frac{m - M + 7.5}{5}.$$

In addition to the Cepheids, lighthouses of the Universe, all stars with a large and known luminosity can be used as milestones in the Galaxy, or indicators of distance: long-period variables (at the brightness maximum) and white stars of a known spectral type. For the former luminosity, like that of the Cepheids, is dependent in a known manner on the variation period; for the latter we can read it from the luminosity-spectrum diagram.

The dependence between period and luminosity of the Cepheids and so-called long-period variables, and also the luminosity-spectrum diagram, were determined using the stars of these types which are relatively close to us and for which the luminosity is known on the basis of reliably determined distances. For nearby stars the distances can be determined directly using the classical method of measurement of trigonometric parallax. For stars that are situated so far away that their parallax is less than the errors in measurement the method is inapplicable and then we must use the method just described.

Everything said above can be applied (and this is particularly valuable to us) to those clusters of stars and to those stellar systems to which the distances are very great in comparison with their dimensions. Then, if in such a distant stellar system there are Cepheids or red long-period variables or white (undoubtedly bright) stars among its members, we can assume that the distance to the system is virtually equal to the distance from us to these members, and we know how to determine their distances. Whether such a star lies at the front edge of the system or at its remote end, at

a great distance, is of no importance. If a group of friends walked far ahead during a stroll in the country and you see that the separation is 2-3 kilometres, it is of no interest to you that this group is spread out along the road through 20 metres. This is the situation with astronomers in their study of a distant stellar system.

Apart from individual stars and stellar clusters, our Galaxy also includes diffuse matter in the form of dark dust nebulae, general layer of cosmic dust, gaseous diffuse and planetary nebulae and the total mass of gas. The latter is mostly invisible neutral hydrogen to be found by its 21-centimetre emission line. The arrangement of diffuse matter should also be studied. Earlier in the book we discussed how distances are worked out to dust and gaseous nebulae.

The most distant objects of the Galaxy, which sort of delineate the contours of its structure and define its size, are the Cepheids, hot giant stars, planetary nebulae, clusters of clouds of neutral hydrogen and globular stellar clusters.

When investigating the motion of the Galaxy's population, astronomers measure "proper motions", i.e. apparent angular shifts (noticeable in pictures for nearby stars only) and radial velocities of very distant objects. For the latter purpose even a giant telescope has to be directed exactly at a faint object for many hours or even nights. The velocities of neutral hydrogen clouds are deduced by examining the profiles of the 21-centimetre line using a radio telescope. Thousands of observations accumulated in the world are corrected taking into account various influences and then subjected to thorough study. From this observational evidence astronomers establish the laws governing the motions, the mass of the Galaxy and density distribution over it produced by the stars per unit volume.

Structure of Our Stellar Home

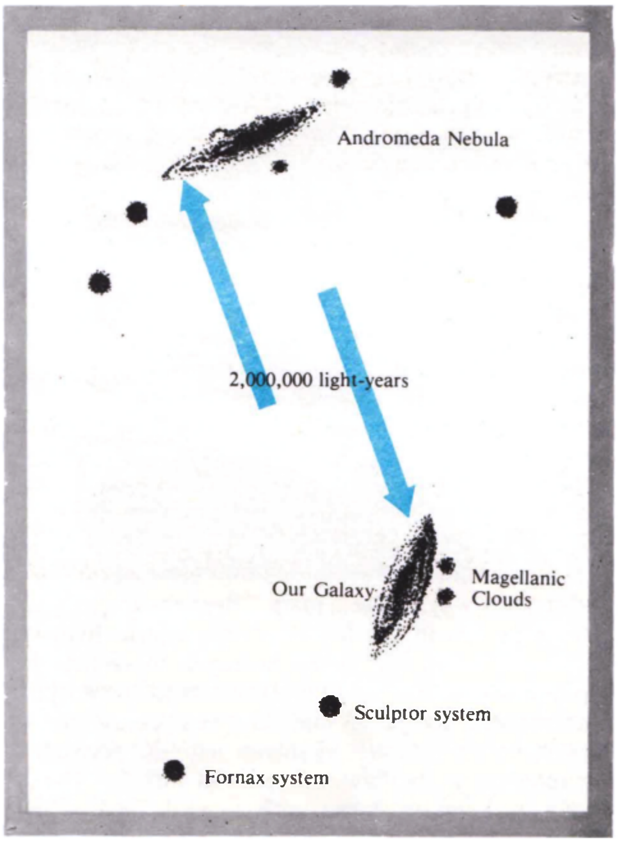
In the final analysis the following was established. Most giant stars and stars of moderate brightness cluster towards the plane of our Galaxy and the centre. The Galaxy has no sharp boundary and its edges thin out gradually. For this reason and also due to inevitable differences in conclusions of various workers, the size of the Galaxy given by them in books published at different times can be somewhat different. But we can assume that the Galaxy is about 100,000 light-years in diameter and from 10 to 15 times smaller in thickness.

Different sorts of populations in the Galaxy are distributed differently. The largest and brightest stars – supergiants – show a tendency to concentrate near the Galaxy's plane. The same is true of a layer of cosmic dust and interstellar gas, whose thickenings appear as nebulae. Near hot stars gas is ionized and we see it as light nebulae, whereas elsewhere in the Galaxy hydrogen is neutral and invisible, accounting for the bulk of the gas. This matter becomes the denser the closer to the galactic plane. Sporadic high-velocity gas clouds occur more than 300 light-years away from the galactic plane.

In the centre of the Galaxy lies the nucleus, which must look like a somewhat flattened ellipsoid of revolution, just like the nuclei of other stellar systems (to be discussed later in the book). The solar system is a bit more than 25,000 light-years away from it. The Galaxy's nucleus contains no hot supergiants and diffuse nebulae whose glow is excited by the former. You will find no dust there either, but there is some neutral hydrogen, which, for some unknown reason, spreads away in the galactic plane at about 50 kilometres per second. The nucleus seems to be surrounded by a fast-rotating ring of neutral hydrogen, and the nucleus's emission seems to be due to K-type orange giants (not supergiants) and a wealth of M-type dwarfs. They are not seen individually and the above assumption follows from analysis of the overall colour and spectrum of the nucleus. Roughly speaking, the Galaxy resembles a lentil or a thin lens in shape, the thicker and brighter nucleus lying in the centre. The nucleus would appear extremely bright, were it not for the mass of cosmic dust that absorbs much of the light. But dust absorbs infrared radiation to a lesser extent and is almost transparent to radio waves. Therefore, observations at these wavelengths bring out the nucleus and its *core*, a small star cluster with a small, but intensive, radio source called Sagittarius A. At present the latter is being studied intensively.

Our position in the Galaxy is extremely unfavourable if we want to study it, since we see it from within. This makes it very difficult to establish what could be found by casting a fleet glance at it from beyond the Galaxy.

We can figure out what our home looks like by examining other homes visible from our window. Here



The Local Group.

our home is the Galaxy, other homes are other galaxies.

Also by analogy with other oblate stellar systems it has long been supposed that our Galaxy in its plane must have spiral arms that begin at the nucleus and wind round it. It was not possible to detect these long spiral arms lying within the main stellar disk before astronomers found a way to measure distances to extremely distant objects and find them in the skies. In other systems the spiral arms, at times very wide and appearing bright against the background of the disk, are best delineated by hot giants, open stellar clusters and diffuse gaseous nebulae. But in our Galaxy we even now cannot, partly due to the disturbing influence of the interstellar dust, study these objects from all aspects. Not infrequently, the spiral arms are not solid and regular, and sometimes they have branches. Therefore, there is no consensus on the arrangement of spirals within the Galaxy. Using radio techniques, astronomers also find spiral arms in regions occupied by neutral hydrogen. But these arms, traced in a larger volume, cannot as yet be readily correlated with the suggestions of arms derived by visual observations. It may well be that these assumptions will turn out to be real.

In recent years much ground has been gained by the theory of very-long-existent spiral arms, which somehow emerged due to running density waves in the galactic disk. At compressions in the arms emerge stars – hot and other young stars that arrange themselves as spiral arms.

Planetary nebulae and novae are interim systems. They moderately concentrate to the galactic plane and markedly to the Galaxy's centre. This system is not plane, it is rather spherical. Almost spherical stellar systems are the few globular clusters markedly concentrated towards the centre. It is these clusters that are to be found up to the remote boundaries of the Galaxy, thus defining its maximal size.

Lastly, a globular system is formed by stars of moderate luminosity – subdwarfs and cluster Cepheids – occurring in quantity in the almost globular nucleus of the Galaxy. The faintly glowing globular crown formed by these stars thus contains the nucleus of the Galaxy and its population I, in which the spiral arms are the brightest. Spiral arms that are conspicuous in other galaxies are an impressive, although light-weight, addition to a near-globular system consisting of faint stars, whose mass, however, is less than that of the supergiants but not to such a degree as their luminosity. The nature of spiral arms – that ornament of some of the galaxies – is as yet unclear.

The Galaxy's mass, as estimated in several ways, amounts to $2 \cdot 10^{11}$ solar masses. Nearly 1/100 of it is interstellar hydrogen, mostly neutral. The above mass figure is derived from the Galaxy's stellar population, hence nonglowing stars, if any, only account for an extremely small proportion of the total stellar mass.

In 1927 Oort found that the stars, including the Sun, revolve about the Galaxy's centre. A way to detect the revolution was indicated by the Russian astronomer Koval'sky as early as last century, but his discovery was forgotten. The Galaxy revolves

not as a wheel, but at the same time not as the planets do around the Sun. It revolves in a complex way, which is a combination of the two above-mentioned patterns of revolution. The solar system revolves at 220 kilometres per second about the Galaxy's centre, which is 25,000 light-years away from us. We do not as yet know the orbit's shape precisely, but if it is almost circular, which is quite probable, then it takes the Sun around 270,000,000 years to make one turn. This period, if you wish, can be viewed as a "space year" suitable for measuring very long periods of time. The entire history of mankind is just a fleeting moment in comparison with this period. If we could see the Sun tearing along and turning in its orbit, just as we see a train turning as it travels along the track, we would not be able to observe the planets revolving around the Sun. They would appear to spin faster than the blades of a fan.

The stars revolve about the Galaxy's centre with different velocities. So cluster variables lag behind the Sun by 100 kilometres per second, and the motion of the solar system at 20 kilometres per second in the direction of the constellation Lyre is its motion within our star cloud or the Local System. It is slow and does not prevent us from revolving round the Galaxy's centre together with the entire Local System.

It can be established that invisible neutral gas occurs throughout the entire volume of the Galaxy. The following is of importance here. Optical studies in the galactic plane are hampered by light absorption by cosmic dust. For radio waves there is practically no absorption, and so the Galaxy is transparent to radio "rays". On the other hand, the Doppler shifts of hydrogen lines with $\lambda = 21$ centimetres, coming from clouds separated differently from us and moving with different velocities, enable these lines to be studied separately. As a result, unlike the nebulae and stars, neutral hydrogen can be studied up to the most remote regions of the Galaxy.

The investigations of gas distribution over the Galaxy have shown that in long condensed arms about 200 parsecs wide the average hydrogen concentration is 1 atom per 1 cubic centimetre, and in between them it is one tenth as high.

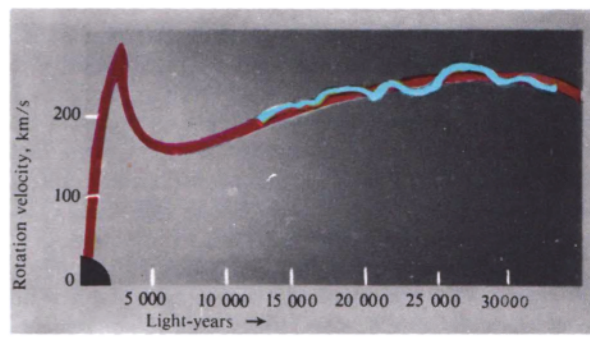
In the centre of the Galaxy, in terms of mass, gas accounts for a negligible proportion of the stellar mass, but in its far reaches gas makes up about 15 per cent, since the stellar space density there drops. On the whole, the mass of gas constitutes around 1 per cent of that of the Galaxy, the balance being the stars. More than 90 per cent of the interstellar hydrogen is neutral. It is only ionized where there are many hot giants, which is generally the case in the middle of spiral galaxies. In our Galaxy the ionized hydrogen reaches 40 per cent at distances from 3,000 to 3,500 parsecs from the centre.

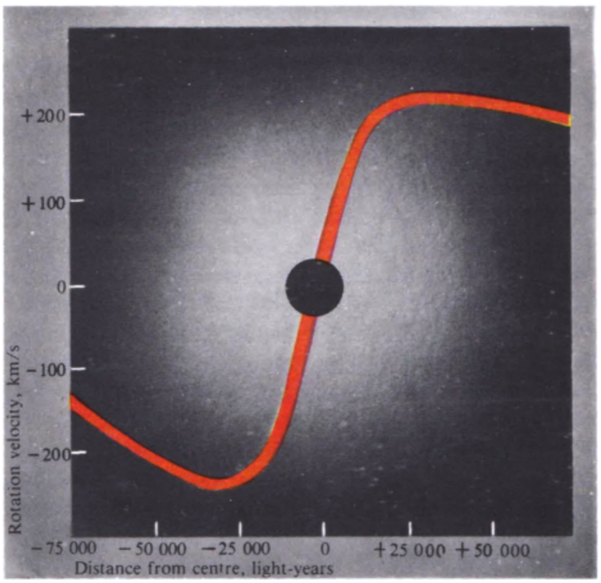
This would suggest that light gaseous diffuse nebulae lie where there are strands of thickened neutral hydrogen. It was also expected that, by analogy with other spiral galaxies, both light nebulae and neutral hydrogen and hot stars, in particular the clusters of the latter, must define the contours of the spiral arms of the Galaxy.

Such correlations are inconvincing since the spiral arms are composed of objects that get together in an arbitrary way. The main discrepancy, strictly speaking, lies in the fact that the neutral hydrogen filaments found are rather arranged in circles than spirals. We believe that the distances to these optical objects are as yet found unreliably, just as the distances to neutral hydrogen clouds derived by extrapolating the law of rotation of the Galaxy.

Spiral galaxies come with spread two or four spiral arms or with nearly concentric arcs. It may well be that our Galaxy is of the latter type: its spiral arms either spread out markedly or consist of an infinite number of short arcs. It would then be clear that the pieces of these formations would hardly be arranged along regular schematic curves,

Rotation velocity in the Milky Way.





Rotation velocity in the Andromeda Nebula.

which are almost nonexistent in real galaxies.

Dutch astronomers have established that in the Galaxy's centre there is a disk with a radius of about 400 parsecs and thickness of about 130 parsecs, its periphery revolving with a velocity of 200 kilometres per second. At a distance of 300 parsecs they have found a ring, or rather a part of a spiral, receding from the centre at 50 kilometres per second. In addition, it has been found that the interstellar gas layer is somewhat skewed relative to the galactic plane, being raised in the direction of the Magellanic Clouds and lowered in the opposite direction. This seems to be due to the influence of these small irregular galaxies (our satellites) on the gaseous layer of our Galaxy. Similar phenomena have been found earlier with some other pairs of galaxies.

From the intensity of the 21-centimetre line the temperature was measured of the interstellar gas in places where it was sufficiently nontransparent and radiating as a black body. The temperature appeared to be 125 K, not 10-15 K, as was believed before. It is assumed now that clouds heat up to 3,000 K in collision, whereupon they cool down to 25 K with the result that different clouds have different temperatures.

To conclude, in the Galaxy the amount of cosmic dust is about one tenth of that of diffuse gas.

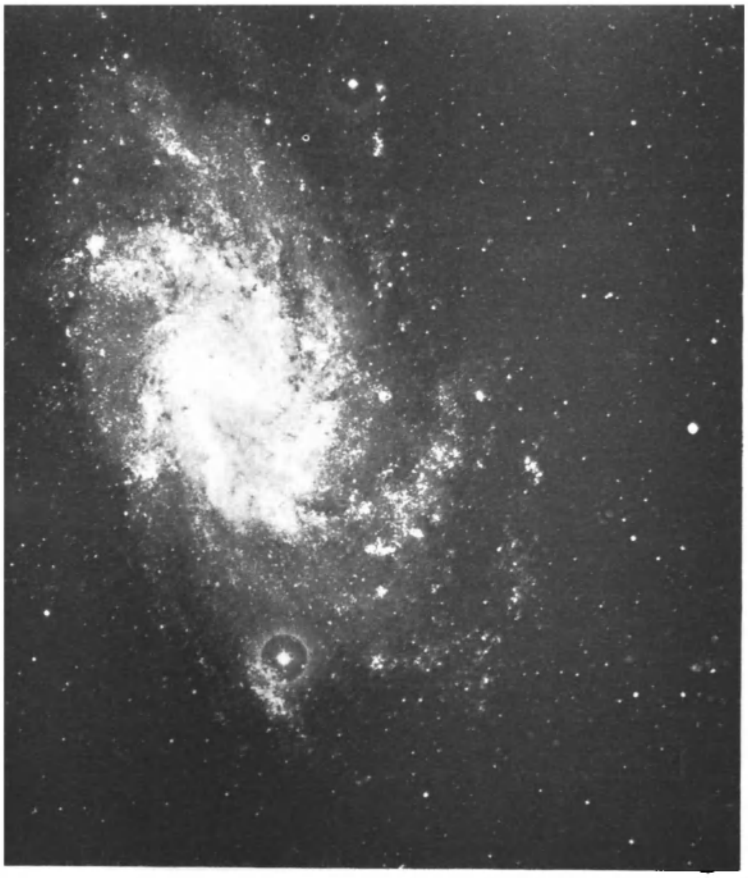
Galaxies – Islands of the Universe

Soon after the invention of the telescope the attention of observers was attracted to a great many light spots of nebular form, which therefore were called nebulae and which were always visible in the same places in different constellations. They were entered into catalogues, but primarily with a feeling of vexation that they hindered the discovery of comets, which have the same appearance as nebulae, but differ by their motion on the stellar skies, like the planets.

The first such catalogue was compiled in the 18th century by the French scientist Messier. In the catalogue, which includes about a hundred of objects, nebulae and stellar clusters are denoted by numerals after the letter M. So, a globular cluster in Hercules is M 13, a large nebula in Andromeda is M 31, and one in Triangulum M 33. Another of such catalogues (NGC) was discussed earlier in the book.

By the use of fast telescopes Herschel and his son John, and later Rosse (also in England), discovered a great many such nebular spots, and by the end of the last century Rosse had discovered a spiral form in certain of them. In such spiral nebulae the nebular core brings forth arms or branches. These arms twist around the nucleus as a spiral and resemble a clock spring. Their nature remained a mystery for a long time until in 1924 Hubble succeeded in obtaining exceptionally sharp photographs of spiral nebulae by using the largest telescope in existence at that time. The edges of these nebulae were found to consist of a great number of extraordinarily faint stars. The nebula was resolved into stars. It became clear that closer toward the centre the continuous nebular glow is produced solely as a result of our perception of a myriad of very close stars, merged into an apparently continuous mass. These photographs revealed immediately that we were not dealing with clouds of dust, shining by reflected light, nor clouds of thin gas, but rather extraordinarily distant stellar systems in which there are incomparably more stars than in globular star clusters.

Those spiral nebulae that have not still been resolved into stars are undoubtedly stellar systems,



The spiral galaxy M 33.

but too distant for their structure to be discerned with present-day telescopes.

In 1944 Baade succeeded in resolving into stars the central part of the spiral nebula M 31 in Andromeda and two small nebulae of elliptic form in its neighbourhood. It was assumed before that elliptic nebulae and the central parts of spiral nebulae consist not of stars, but of gas or cosmic dust.

Spectra confirm the stellar nature of the centres of spiral nebulae. These are absorption spectra, very similar to the solar spectrum, showing that the majority of stars in the spiral nebulae are yellow stars of the solar type. The spiral arms consist of hotter white stars. From the displacement of the dark lines in the spectra of spiral nebulae it was possible to determine the velocities of their motion. As a whole, they move with velocities of hundreds of kilometres per second.

Finally the nature of spiral nebulae was revealed when Cepheids were found in them on the photographs. There were also long-period variables and bright bluish stars. Globular star clusters, completely identical to clusters in our Galaxy, later were discovered in the spiral nebula in Andromeda, but as a result of their great distance they were scarcely distinguishable in appearance from bright stars. Also discovered in the spiral nebulae were immense clouds of thin gas, giving a spectrum of bright lines, again similar to the one encountered in interstellar space within Galaxy. It was found that in the globular star clusters, and also in the elliptic star systems, the stars form a different spectrum-luminosity diagram from the one which applies to the stars making up spiral arms and the irregular stellar systems of the Magellanic Clouds type. They are seen by the naked eye and resemble scraps of the Milky Way.

Distances, and hence dimensions, are determined the most accurately from the apparent brightness of the Cepheid variables when the latter are observed within a given galaxy. This is only possible for the nearby galaxies. For more distant galaxies distances are worked out from the apparent brightness of the brightest supergiants in them. In elliptic galaxies, which in outward appearance resemble globular clusters within our Galaxy, only larger in size, there are no supergiants.

The most interesting among the elliptic galaxies is the brightest and largest M 87, the main feature of the galactic cluster in Virgo. This giant galaxy has an entourage of hundreds of globular stellar clusters, which are so distant that are just barely distinguishable from stars in photographs. On the other hand, elliptic galaxies – companions of the large spiral in Andromeda (M 31) – are far smaller than M 87. A recent discovery is dwarf elliptic galaxies, which are only several times larger and brighter than a typical globular cluster.

Most of galaxies are so far away that separate stars are indistinguishable in them. Therefore, the above-mentioned ways of determining distances to them are of no avail. At the same time, luminosities and linear dimensions of galaxies are so varied that neither their apparent angular diameter nor their apparent brightness can be a measure of their distance. Their distances are estimated based on an amazing property of all the galaxies discovered by Hubble.

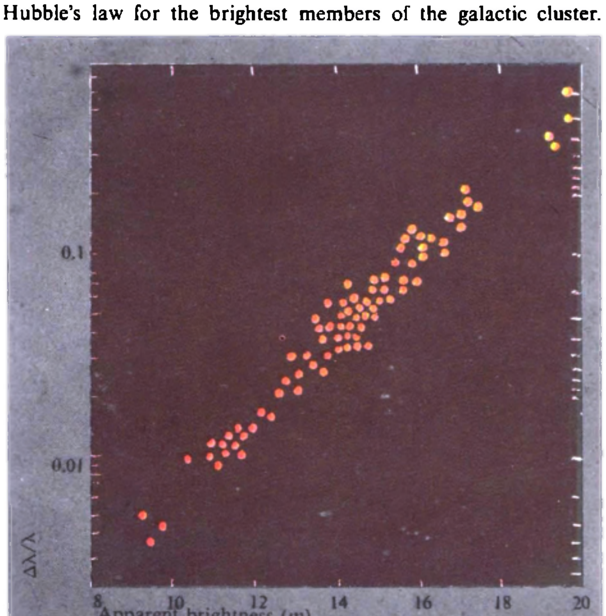
Studies of galaxies with known distances and velocities of motion along the line of sight indicated that their spectral lines are shifted towards the red end of the spectrum by an amount proportional to their distance. The velocity at which a galaxy recedes combines with the proper velocity of the galaxy,

which is generally not higher than several hundreds of kilometres per second. For nearby galaxies the proper velocity and that causing the red shift are of the same order of magnitude, but for distant galaxies the velocity giving the red shift is much larger – thousands and tens of thousands of kilometres per second. It is for this reason that the distance to the distant galaxies is determined from the red shift with the least relative error. For example, in a close group of galaxies individual members move at 200-400 kilometres per second relative to one another, whereas the group as a whole may have an average red shift corresponding to 1,000 or 10,000 kilometres per second. According to modern estimates, "Hubble's constant", the increment of red shift each 3,000,000 light-years (or 1,000,000 parsecs), is about 100 kilometres per second. In this example in the first case the galaxies would be around 10,000,000 parsecs away, the possible error being no more than 20-30 per cent. In the second case the distance would be 100,000,000 parsecs with an error not higher than 3 per cent.

The spiral nebula in Andromeda appears larger and brighter than all the others because it is the closest to our Galaxy (yet nearer are only the Magellanic Clouds). The distance between the Earth and this nebula is more than two million light-years – and this is the closest one! Its light coming to us now left the nebula when man had not yet appeared on Earth. It measures 100,000 light-years in diameter, but in the direction perpendicular to the plane of its maximum extent it is many times thinner, being much flattened. Comparing the appearance of the nebulae such as in Triangulum (almost circular outer outlines), in Andromeda (oblong or elliptic) and in Virgo (spindle-shaped), we must conclude that the difference in their appearance is caused by their foreshortening relative to us. These stellar systems (which, as we now know, have the complete right to be called galaxies, since they are enormous stellar systems like our Galaxy), obviously have a flattened lens-like form and frequently a spiral structure. The galaxy in the nebula in Triangulum lies before us "edge-on", the galaxy in Andromeda has its plane of symmetry tilted toward us, and the galaxy in Virgo has its side turned toward us. It is worth mentioning that along the spindle it is possible to see a dark band. Such dark bands can be seen in many galaxies having a spindle-like form. It is quite probable, as pointed out by Curtis (United States), that this is a cluster of dark nebulae consisting of dust and concentrated toward their equatorial planes. In other galaxies, less tilted in our direction, it is also possible to note dark regions against the background of the shining mass of the nucleus, in the arms and between the arms of the spiral curls. There is absorbing matter in all galaxies, and not only in those turned to us edge-on. This is an additional similarity between distant galaxies and our Galaxy.

By putting on the spectrograph slit the different parts of the images of galaxies obtained through the telescope it was possible to measure their radial velocity. It was found that galaxies rotate around their short axis, perpendicular to their equatorial planes. The spiral galaxy in Andromeda, in its inner parts, rotates as a solid body, like a cart-wheel. This means that its inner parts, which provide little light and apparently therefore contain few stars, nevertheless have a great mass. In the galaxy M 33 in Triangulum, the innermost parts, within 3,000 light-years from the centre, also rotate as a solid body. Outwards, on the other hand, the rotational velocity decreases very

Hubble's law for the brightest members of the galactic cluster.



rapidly. It therefore follows that, as in the galaxy in Andromeda, the bulk of its mass is concentrated in the central region of the stellar system. This mass has been estimated at a hundred milliard solar masses from the observed rotational velocity.

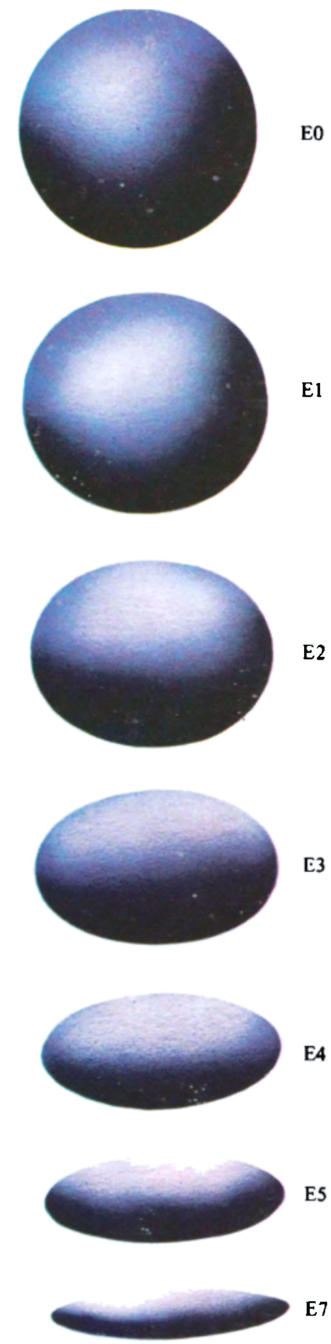
Ambartsumyan has computed the brightness of that region of our Galaxy in which the solar system is situated as it would appear if we could look at the Galaxy from afar, just like we see other stellar systems. Comparing this estimate with the brightness of various parts of the galaxy in Andromeda, he arrived at an unexpected conclusion.

In the galaxy in Andromeda, the stellar space density corresponding to the density of stars in our Galaxy in the neighbourhood of the Sun is at a distance of 5,000 light-years from its centre, a place where on the photographs it is possible to see only the scarcely visible edges of this galaxy. Consequently, he postulates by analogy that we live on the most distant outskirts of our stellar system, where the population is very thin. We are probably situated beyond the limits of the spiral arms, where the stellar density is smaller.

As we see, stars are grouped in space in gigantic systems, frequently of a spiral form. They are scattered, like islands, in the ocean of the Universe. Islands of the Universe or island universes – these are names often given to galaxies. In certain places, such as in the constellation Virgo, the galaxies are grouped into clouds of galaxies – the islands of the Universe form an archipelago. The clouds of galaxies or groups of islands of the Universe resemble open star clusters, but on a greater scale.

Initially astronomers were perplexed by the great difference between the dimensions of our Galaxy and other galaxies. The history of science has taught astronomers to be humble, which their predecessors were not, assuming that the Earth was the centre of the Universe and that it occupied a special position in the Universe.

As we have seen, light absorption in the Galaxy was discovered finally in the 1930s. An allowance for its influence on the apparent brightness of stars led to a considerable decrease in the dimensions of the Galaxy. On the other hand, the distances, and hence the dimensions, of other galaxies appeared to be somewhat larger than was thought earlier, since a careful treatment of photographs revealed faintly luminescent outer parts of galaxies that had not been



The classification of elliptic systems.

noticed earlier. As a result, the dimensions of our Galaxy and others were found to be less different from one another.

Undoubtedly, when it became possible to investigate the more distant galaxies in more detail we found some greater than ours. But in the long run, having learned that the Earth is not the centre of the Universe, that it is not the largest of the planets, and that our Sun is not the largest, nor the brightest, after all these blows at our false pride we cannot permit ourselves the luxury of assuming that we live in one of the largest galaxies, although near its edge. We are inhabitants of a far "outbuilding", although it is one of the largest "houses" in a stellar city known as the Metagalaxy.

More About Galaxies

Galactic studies are at present the most exciting domain of astronomy, since it is this field that is the richest in developments and is leading us to revealing the most general and shattering properties of the Universe. The following sections are therefore devoted to galaxies.

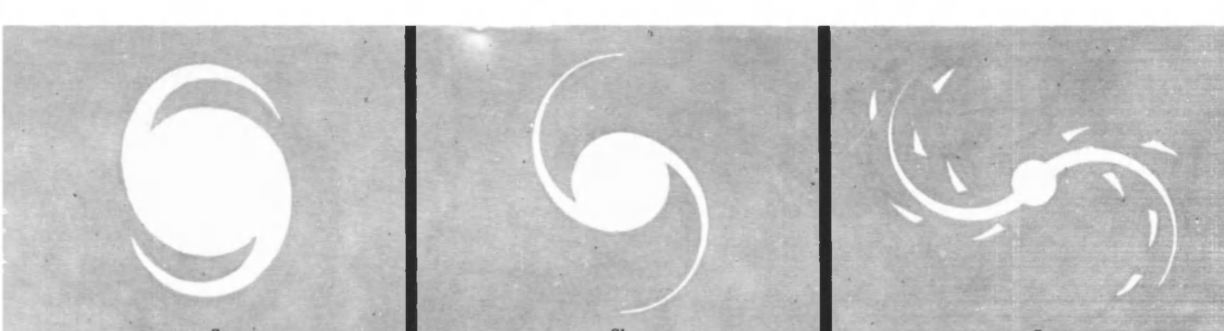
In the 1920s Hubble compiled a first, simple classification of galaxies, whose deficiency became to be felt only recently after much larger population of these celestial objects have become known to us.

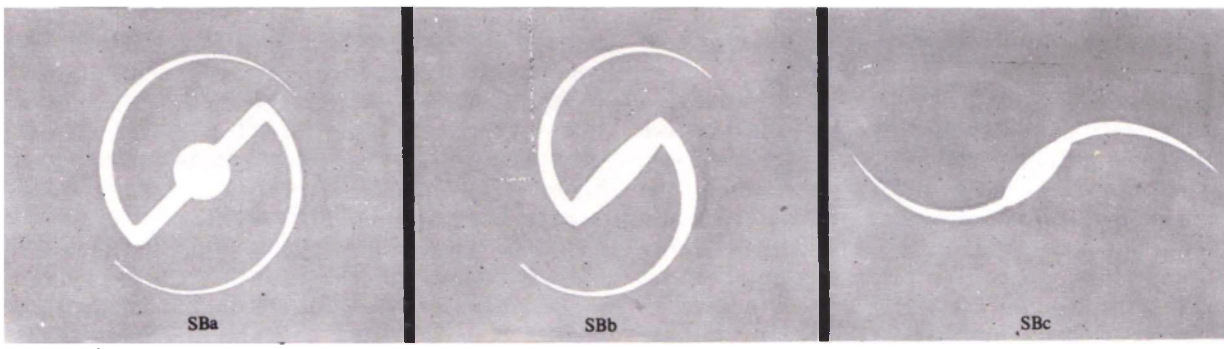
Hubble singled out elliptic galaxies, denoted by E, which superficially resemble the globular clusters within our Galaxy, but are vaster. They are irregular, contain no hot stars, supergiants, dust and gaseous nebulae. Their stellar space density falls off slowly and smoothly with distance from the centre, which contains no nucleus. Such galaxies are legion. An example of them is M 87.

Next he distinguished "irregular" galaxies, denoted as Ir, of ragged structure and irregular form. They are smaller than elliptic ones and fewer in number. Their brightness and luminosity are smaller, they are oblated strongly but rich in hot supergiants, gaseous nebulae and dust. Examples of them are the Magellanic Clouds, both Large and Small.

Hubble subdivided spiral galaxies into two families: normal S and "barred" SB. In the former, arms originate directly from the nucleus; in the latter, the nucleus is crossed by a wide, bright strap, called a bar. The spiral arms originate from the bar's ends. Besides, sometimes a light ring passes through the bar's ends. For both species of spiral galaxies Hubble established three types, denoted by the letters a, b, and c. In Sa and SBa galaxies the nucleus is the brightest and largest, the arms being weak, structureless and amorphous. In Sb and SBb galaxies arms are brighter and somewhat ragged, and the nucleus is relatively smaller and fainter. M 31 is an Sb galaxy. Sc and SBc galaxies have a small and faint nucleus, their arms are strong, bright, markedly ragged. M 33 in Triangulum is an Sc Galaxy. The arms increase in their raggedness due to an increase in the number of hot giants and their groups, bright gaseous nebulae, open clusters, and in Sc type also of superassociations. Hot giants are responsible for the arms being bluer than the nucleus, they become bluer in passing over from Sa to Sc type. Among elliptic and spiral galaxies the largest ones have the same luminosity and size. Their absolute stellar magnitude is -21^m, thus suggesting that they are thousands of millions of times brighter than the Sun.

The classification of spiral nebulae.





The classification of barred spirals.

Also of importance is the ratio $M : L$. With elliptic galaxies it is of the order of tens, smaller than with spirals; with irregular galaxies it drops to about 2-5. This is due to a larger contribution of supergiants to the total luminosity of the system, as the latter grows much faster than the mass.

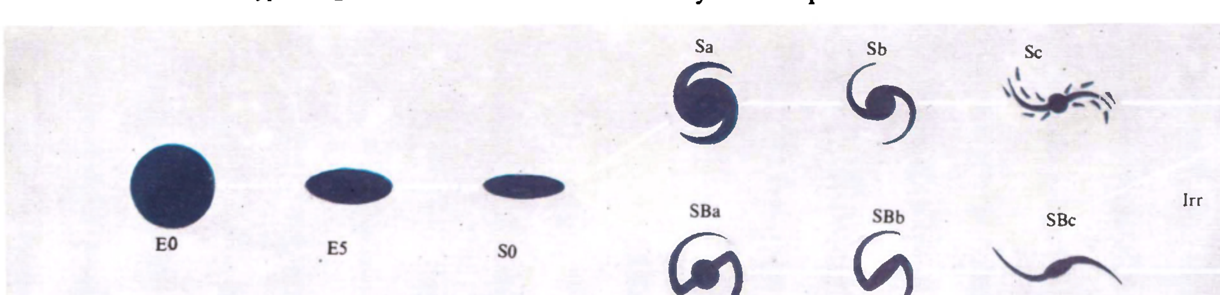
Spiral and irregular galaxies produce moderate radio emission, just like our Galaxy, and it is caused by the same reasons.

It has been found in recent years that the galaxies have a far more variegated nature than was assumed by Hubble, although he himself later introduced S0 and SB0 types, intermediate ones between E and spirals. These are characterized by the presence of a flat disk (population I) around a large and bright nucleus, but in that disk there are no dust and gases and no spiral arms.

First of all, it appeared that in our neighbourhood there are several extremely faint dwarf galaxies. Some of them are irregular, others are globular, but so tenuous that in photographs they appear as hardly discernible spots, although not that small. And fairly recently a binary galaxy has been discovered – a pair of pygmies just a bit brighter than the brightest of globular clusters. The pygmies differ from these by their independent position in space and the presence of a mass of glowing gases, which do not occur in globular clusters. So, specifically, elliptic and globular galaxies vary in luminosity, mass and concentration of stars from supergiants to those similar to globular clusters, and from exceedingly tenuous and transparent to very compact and highly concentrated. These compact galaxies, discovered by Zwicky in 1964, include the above-mentioned pygmies. In the photographs obtained using the largest telescopes, compact galaxies are just distinguishable from stars. Sometimes they can only be differentiated by their red shift in their spectrum, many among them having high luminosity.

Zwicky also found that the galaxies that similarly appear as elliptic ones if central parts are overexposed in photographs, differ only in that some of them have a tiny star-shaped nucleus in the centre, and others

Hubble's classification of types of galaxies.



have none. The author has also found quite a number of galaxies, that completely fall out of Hubble's classifications or its modifications. The most interesting of these are the galaxies with oppositely twisted arms, and the numerous ring galaxies having both amorphous and ragged structure. There are galaxies with dust population I, but without a bright component either in the form of a disk or in the form of spiral arms. Also there are galaxies with compound nuclei surrounded by a colossal halo. At an earlier date galaxies have been found that are irregular although not ragged but amorphous, i.e. free of hot stars and their clusters. They are denoted by Ir II.

With many galaxies the author has found internal and external arms of absolutely different structure (amorphous and ragged), interweavings and crossings of the arms, figure-of-eight arms that turn into circles or loops. These shapes cannot be attributed to some mechanical processes and they resemble perturbed lines of force of the magnetic field of a sphere.

By and large, the world of galaxies appeared to be amazingly diverse. Some other examples of this variety will appear in later sections.

The author and his co-workers at Moscow University have recently issued a catalogue containing 30,000 galaxies and giving their position in the skies, brightness, dimensions, colour, velocity, detailed description and references to evidence available for each of them. This catalogue is denoted by MCG (Morphological Catalogue of Galaxies). It covers all the galaxies brighter than 15^m to a declination of $-45°$.

Now let us dwell on some of the nearby islands of the Universe in more detail.

The Magellanic Clouds in the constellation Dorado, companions of our Galaxy, are of immense interest to us in that they are the nearest galaxies (about 50,000 parsecs away), whose structure and motion, as well as the brightest objects in them, lend themselves to the most detailed examination. So, for example, in them we can measure brightness and colour of stars brighter than $+1$, whereas those brighter than -6 can be investigated in some detail by spectral means. The major axis of the Large Magellanic Cloud (LMC) is 12 kiloparsecs long, and that of the Small Magellanic Cloud (SMC) is 4 kiloparsecs long. They are surrounded by a common 3×15 kiloparsec envelope of tenuous neutral hydrogen. The two clouds are

NGC 205 in the neighbourhood of Andromeda Nebula, an elliptical stellar system (2.3 million light-years away, 8,000 light-years across).

immersed into it, thereby suggesting that they are not only close to each other, as they appear in the sky, but are also connected by more close ties. This is supported by the fact that a relatively dense gaseous isthmus has been found between the clouds. Their velocities relative to the Galaxy's centre are $+40$ (LMC) and -15 (SMC) kilometres per second. The LMC's mass worked out from its rotation amounts to 10^{10} solar masses, i.e. is one fifteenth of the Galaxy's mass. Viewed as irregular in shape, they, especially LMC, possess distinctive features of barred spirals. Astronomers now study the distribution in LMC of over 200,000 stars with an absolute magnitude of over 0^m.

In LMC are observed Cepheids, cluster and other variables, all sorts of blue giants, gaseous and dust clouds. By the way, there we can find the brightest of known stars – S Doradus. This is a slightly variable star about a million times brighter than the Sun.

In the Magellanic Clouds dozens of open and globular clusters are observed and examined, among these such clusters that have no parallel in the Galaxy not only in terms of size, but also in terms of structure and stellar composition. The Magellanic variables appeared to differ in luminosity from the variables of the same period known in the Galaxy. In other words, after the earliest important discoveries suggesting the similarity of populations of spiral arms of the Galaxy

The spiral nebula M 51 in Canes Venatici (8 million light-years away) is a galaxy similar to ours with the only exception of a small companion galaxy.

and the LMC, further, more detailed studies, brought out minor departures only. But these departures, which seem to be characteristic of galaxies in general, are suggestive of the rich variety of their nature (which is of importance in principle) and make accurate determination of distances to seemingly similar galaxies very difficult.

In terms of per cent relative to the total mass, the Magellanic Clouds possess more neutral hydrogen than any of the known galaxies: from 20 to 30 per cent. Both in the Clouds and the Galaxy the glowing gaseous nebulae appeared to have the same chemical composition.

Irregular galaxies have moderate and low luminosities, they are mostly dwarfs with absolute magnitudes -14^m and diameters from 1,500 to 3,000 parsecs.

The Magellanic Clouds are some of the brightest and largest irregular galaxies.

M 31 is the nearest spiral giant galaxy, and it is thought of as very similar to our Galaxy. But it is 12 times farther away from us than the Magellanic Clouds and therefore similar objects in it appear 1/150th as bright. Examination of the galaxy is hampered by inclination of its plane to the line of sight. Special-purpose explorations revealed a multitude of Cepheid variables and other bright variables in it; all in all about 170 new stars have been registered, more than in our Galaxy, where we only see the nearby stars. Also several hundreds of diffuse gaseous nebulae have been found in it, which with remarkable regularity, like beads on a string, delineate bright spiral arms.

In recent years it has been found that filaments of cosmic dust accompanying bright external spiral arms are to be traced farther to the centre inside a structureless, seemingly amorphous "lens", or the main body of the galaxy. This enabled some scientists to suppose that spiral arms originate very close to the nucleus in the form of dark, dust arms that then become lighter. It would be, however, more pertinent to say that these dark filaments, at first in disarray and in no way connected with one another, farther away from the centre become thicker and then begin to accompany the bright spiral arms made up of stars. Initially, the arms are amorphous in appearance and contain no supergiants. Then supergiants emerge and their density is the higher the farther away they are from the lens. Also growing is the density of light gaseous nebulae, so that eventually the arms thin out, their spiral structure sort of smears out, although regions of declining stellar space density stretch much farther. Comparison shows that the solar system, if placed at the same distance from the centre of M 31 as it is from the Galaxy's centre, would lie on the boundary of, for the time being, clearly seen spiral arms, in the domain of relatively meagre stellar space density.

M 31 has a small nucleus that rotates exceedingly fast. The nucleus looks like a star and only through the largest telescopes can it be distinguished from single stars, although it is a whole stellar cluster, strongly condensed. Its apparent magnitude is $14^m.5$, and the absolute magnitude is -10^m, i.e. it is somewhat brighter than the brightest globular clusters that belong to the same galaxy. But here the difference in stellar composition is larger: the nucleus of M 31 seems to consist of red and yellow giants of conventional chemical composition, and the globular

clusters include giants that are poorer in metals.

Another of the nearby spiral galaxies (M 33 in Triangulum, type Sc) has a luminosity that is one sixth of M 31, and a diameter that is one third of the Galaxy's.

Elliptic galaxies likewise have giants and dwarfs. The brightest and largest of the known galaxies are two elliptic galaxies in the cluster Virgo: NGC 4486 (M 87) and NGC 4472 (M 49) with diameters 22,000 and 31,000 parsecs, respectively. The boundaries of elliptic galaxies are yet more arbitrary than those of spiral galaxies. If we take the boundary to be a place where the surface brightness is barely distinguishable from the background of a clear night sky, then the size of supergiant elliptic and spiral galaxies will appear to be about the same and, as we have seen, will be around 30,000 parsecs, or almost 100,000 light-years. But in the neighbourhood there are no giant elliptic galaxies, and their distances, and hence their luminosities and dimensions, are determined from red shift.

In our neighbourhood are only dwarf elliptic galaxies, companions of the spiral galaxy M 31 in Andromeda, their absolute magnitudes being about -15^m, and their size being about 3,500 parsecs. "Extreme dwarfs" are very faint: the absolute magnitude ranges from -12 to -8, they are only 3,000 light-years across. Belonging to this class are the faint spheroidal galaxies in our neighbourhood – in Fornax and Sculptor – and even more so the systems Leo I and Leo II.

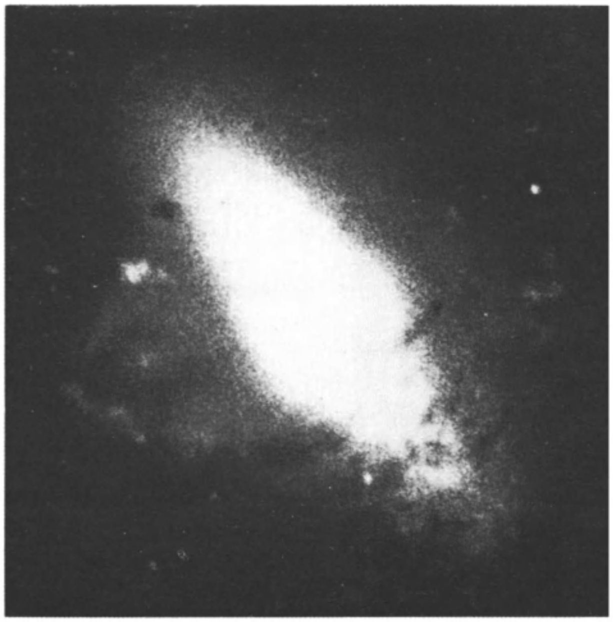
The total luminosity function for galaxies is, generally speaking, not yet established, and only for the brightest galaxies in clusters it is known in a more or less reliable way. It is assumed that it may differ from cluster to cluster. This question is very difficult to answer due to the fact that it is impossible to single out in all certainty the galaxies belonging to the cluster from those that happen to project on it.

The Galaxy lies within an isolated group called the *Local Group* or *Local System of galaxies.* Two main groups with supergiants in each stand out in it. These are our Galaxy with its companions, the Magellanic Clouds and M 31 with its several elliptic companions. After the extreme dwarfs – spheroidal galaxies of the Sculptor type and other irregular galaxies – have been discovered, it turned out that dwarfs predominate in the Local System. There are two supergiant systems in each spiral of moderate size (M 33 in Triangulum), two compact dwarf elliptic galaxies (NGC 205 and 221), two fairly tenuous (NGC 147 and 185), six extremely tenuous spheroidal (in Fornax, Sculptor, Leo I, Leo II, Ursa Minor, Draco), irregular galaxies (Magellanic Clouds, NGC 6822, IC 1613, the Wolf-Lundmark system, three Holmberg's systems, and maybe three dwarfs in the Sextans, all of which are as yet little studied). So, our Local System has two giant spirals, one medium-sized spiral and from 17 to 20 dwarfs, mostly elliptic and spheroidal. It appeared that dwarfs dominate, and the average absolute size of galaxies is now strongly shifted towards low luminosities.

It is not clear to what extent the luminosity function of the galaxies in our neighbourhood can be ascribed to galactic clusters and the entire Metagalaxy. Clusters are dominated by elliptic galaxies and it is these that are often the brightest, but in our neighbourhood there are no elliptic supergiants. The author has found that there are groups of large galaxies without dwarf companions.

The density of dwarfs in the general field of the Metagalaxy and galactic clusters can, therefore, differ from that in the Local System. Faint dwarfs, hardly discernible even in the neighbourhood, cannot be seen

The peculiar galaxy NGC 2685 (a radio source).



at large distances. But astronomers seek brighter dwarfs in the nearby clusters. In 1956 about 50 dwarfs with brightness only slightly concentrated to their centres were found in a cluster in Virgo. Their absolute magnitude is about -13^m. But they are thought to be different from the Sculptor type galaxies, which are two or three units fainter. They are referred to a new type, IC 3475. Three years later in a cluster in Fornax, too, 16 dwarfs were found that feature weak concentration of light to the centre. Thus, large open clusters include dwarfs, but their percentage seems to be lower than in the Local System.

Zwicky believes that the luminosity function of galaxies must continue down to small systems, such as globular stellar clusters in the Galaxy and even fainter, but his opinion is not shared by other workers.

All the above also applies to the statistics of galactic types. According to de Vaucouleurs, among about 1,500 bright galaxies elliptic ones account for 13 per cent, ones that are normally referred to as transitional (S0 type) account for 21.5 per cent, spiral ones account for 61.1 per cent, irregular ones account for 4.4 per cent. The most numerous spiral galaxies are dominated by Sb, Sc, and SBb. "Early" Sa spirals are rare, these are galaxies of great apparent brightness, more than the 13th apparent magnitude. Our neighbourhood and nearby clusters are dominated by elliptic galaxies.

We have already mentioned that spiral galaxies observed edge-on show an equatorial layer of cosmic dust, which appears as a dark band. The author has found that its thickness differs from galaxy to galaxy. Dark matter can also be "felt" in spiral galaxies seen edge-on in the form of individual spots in spiral arms running along their inner or outer side or in the form of branched canals (as in M 33). Sometimes the dust layer is disposed along the bar.

Complexes of light diffuse nebulae are directly observed even in fairly distant spiral and irregular galaxies having many hot stars and their clusters. These galaxies are very ragged. Some of the nearby galaxies, such as M 33 in Triangulum, M 31 in Andromeda, the Magellanic Clouds, even contain individual diffuse nebulae.

The Large Magellanic Cloud has a giant complex of gaseous nebulae that envelops a vast cluster of hot giants. The complex is called the Tarantula Nebula and the cluster is called a superassociation. If the Tarantula were at the place of Orion, its light would produce shadows on Earth.

It has been found that M 31 contains several hundred diffuse nebulae, which are so arranged that they delineate the spiral arms but not completely coincide with the stellar arms. Being small when observed at enormous distances, they are more conveniently found in the pictures taken in the red hydrogen line H_α (through a red light filter). Such emission concentrations are used to investigate into the revolution of the outermost parts of galaxies whose stellar spectrum is too faint to be recorded.

Integral spectra of many galaxies show H_α and λ 3727-29 Å lines (forbidden lines of ionized hydrogen) produced by the total light of component nebulae. When the bright H_α line is seen throughout a galaxy, it is used to study the rotation of the system.

It has been shown statistically that the earlier are galaxies, i. e. the fewer are hot giants in them, the rarer are bright lines in their spectra. Elliptic galaxies are essentially gas-free. Several galaxies have been discovered in which bright lines are even stronger than in irregular galaxies.

Radio observations already enable astronomers to detect the thermal radiation of gases in nearby galaxies and even work out their velocities at various places, their rotation being established from the 21-centimetre line. For these stellar systems first neutral hydrogen distribution maps began to be charted.

The total mass of gas in galaxies of different types has been estimated to account for the following percentage of their total mass:

Irregular	Sc	Sb
17	8	1

(our Galaxy belongs to Sb or Sc).

In the M 31 and Magellanic Clouds planetary nebulae have been found, their luminosity being as high as -3^m.

Hydrogen-to-helium ratio in the nebulae of other galaxies was estimated to be the same as in our Galaxy, thus suggesting that the chemical elements have the same abundances in the Metagalaxy.

Diffuse matter is thus of immense importance in the cosmos.

Tailor-made artificial satellites found in 1975 nearly 200 X-ray sources in space. These are mostly envelopes cast off by supernovae, neutron stars, or pulsars, which

are believed to have been formed during the explosions of supernovae; and companions of some of the stars. Plasma in some of the white dwarfs seems to be able to produce X-ray radiation as well. Our Sun is also an X-ray emitter.

In addition, there are many extragalactic X-ray sources. These are some radio galaxies, their nuclei, and some extended sources associated with the plasma that is spread throughout galactic clusters.

Galactic Groups and Clusters

Many binary galaxies are known that are similar in brightness and size. Also there exist dwarf galaxies, companions. So M 31 has two close companions, dwarf elliptic galaxies.

Unfortunately, we do not know the distances of most of the galaxies, which are arranged close to one another in the skies. Therefore, it is normally unknown whether a given faint galaxy is really a companion or a more distant galaxy that just happened to project near a bright one.

Statistical calculations indicate that most of seemingly binary and multiple galaxies are such in reality. Multiple galaxies, i.e. small galactic groups, seem to be especially frequent.

An example of a fairly spread group of a wide variety of galaxies is our Local System. There are more close groups as well.

Components of binary and multiple stars are normally hundreds and thousands their diameters apart, whereas members of galactic groups are only several diameters apart. Cases are known where they are adjacent to and partially penetrate into one another.

Multiple stars mainly have the following structure: one close pair of stars revolves at a large distance about a general centre of mass, where there is a single star or another close pair.

Multiple stars with companions spaced by comparable distances are extremely rare. Ambartsumyan called them trapezia, since this type includes a so-called quadruple star, in which each companion lies at a vortex of a trapezium.

Ambartsumyan showed that trapezium systems of bodies are inherently unstable, they must separate soon and their companions must recede from one another and lose the connections. This occurs because their mutual pulls are comparable, act in different directions and in different planes, thus making any sort of stable motion and stable orbit impossible.

Moreover, Ambartsumyan paid attention to the fact that whereas the multiple stars are dominated by systems with stable configurations and orbit types, the multiple galaxies are dominated by unstable configurations and orbits. He concluded that the systems of stars and galaxies that form trapezia are younger formations (because we see them when they have not yet decomposed). Such multiple galaxies must be much younger than our Galaxy, which is estimated to be 10^{10} or more years old.

Galactic clouds and clusters are far simpler to identify. Galactic clouds are diffuse concentrations of galaxies, and clusters are more compact formations, although they too are subdivided into concentrated and diffuse ones.

The most conspicuous are the nearby formations consisting of galaxies that appear brighter and larger. These are neighbouring clouds in the constellations Ursa Major and Canes Venatici, the open cluster in Virgo and dense clusters in Coma Berenices and Corona Borealis. These latter display spherical symmetry and consist mostly of E and S0 galaxies.

The clouds in Ursa Major and Canes Venatici occupy in the skies a vast area of $30 \times 40°$, a cluster in Virgo, no less than $25 \times 40°$, a cluster in Coma Berenices has a diameter of 12°. In clusters and clouds, showing no spherical symmetry, subsystems are noticeable. So in Virgo two fairly close clusters project one upon the other, one consisting mostly of spiral galaxies, the other mostly of elliptic ones. The latter largely make up dense clusters.

The clouds in Ursa Major and Virgo, the nearest to us, are around 10 megaparsecs away. In the former, there are more than 200 members of the 13th apparent magnitude, in the latter more than 150, whereas Coma Berenices has more than 30,000 stars brighter than 19th magnitude. But they must include many, perhaps even much more, dwarf galaxies. The brightest of them have recently been found in the nearby clusters in Virgo and Fornax.

Galaxies tend to concentrate to the centre of a cluster, where frequently the brightest and largest of its members lie (galaxies are distributed along the radius of dense clusters in the same way as do molecules in the so-called isothermal gaseous

Coma cluster, a galactic cluster, about 350 million light-years away.

sphere, thus suggesting that such a cluster is stationary).

In recent years statistical studies of clusters have been under way (as to detailed examination of them, it awaits further improvements in the techniques).

Very rude estimates are available of the number of bright members, their distances, size and position in the skies for nearly 2,000 clusters entered in the cluster catalogue. So in one (though the richest) region, $6 \times 6°$ in size, Zwicky has counted 120,000 galaxies that belong to 100 clusters and are from 200 to over 600 million parsecs away.

To determine distances to clusters reliably requires much tedious work in the form both of night observations and later examination of the pictures. This is because to determine a distance from red shift, spectra of several galaxies in a distant cluster have to be photographed, which requires many hours of laborious spectral photographing. When it is required to determine the distance of a galaxy from its magnitude, its brightness must be correctly measured and several corrections introduced, which are not always known adequately. It should, for example, be taken into account that owing to the red shift of the entire spectrum a distant galaxy appears to be redder, and hence fainter, than in reality. It is then supposed that, say, the 5th (in increasing order of brightness) galaxies in all the clusters have the same luminosity. The validity of this assumption is put in question, and rightly so.

It has long been decided that most of galaxies lie within clusters and that in the "general field" in between galaxies are rarer, thus giving rise to uneven distribution of density within the Metagalaxy. But are there clusters of clusters? Or the clusters fill evenly the visible part of the Metagalaxy?

De Vaucouleurs put forward convincing evidence that bright nearby galaxies form an oblated supersystem, a "supergalaxy". At its centre lies the Virgo cluster, which acts as the nucleus. It includes the Ursa Major cloud and the Local Group of galaxies lying near the symmetry plane of this supersystem. Therefore, the bright galaxies appear for us as arranged in a ring, just like the Milky Way. Many authorities, however, believe that there are no superclusters including many clusters of galaxies. At the same time, Zwicky maintains that the clusters of galaxies are distributed in space fairly evenly, both in our neighbourhood and elsewhere. He also considers that in the intergalactic space there are many star clusters and individual stars, as well as dust clouds. Rich clusters, he thinks, have much dust and it shields from us the more distant clusters.

Some confusion in the treatment of distribution of clusters in space is a serious hindrance for any attempts to compare with reality the various cosmological models: a finite and infinite Universe with Euclidean and non-Euclidean space, expanding or stationary, and so forth.

The world's largest telescope can now "take in" thousands of millions of galaxies – an accurate count would be exceedingly difficult, and hardly necessary. What is important here is that we do not witness any reduction in the number of galaxies and their clusters with distance. In other words, there are no signs that we reached the boundary of the Metagalaxy, that prodigious archipelago of island universes, to which all the visible galaxies belong. So far we have only been dealing with a part of the Metagalaxy. Maybe there are other Metagalaxies as well.

Your Address in the Infinite Universe

We will now summarize the results of the development of our knowledge concerning the place of man in the Universe, insofar as we know its structure at the

present time. We will present this result in the form of the reader's address. We have written your address for you, and you need only complete the first line.

1. Your country, town, etc.
2. Planet Earth,
3. Our Solar System,
4. Star Cloud "Local System",
5. Our Galaxy,
6. Local Cluster of Galaxies,
7. Our Metagalaxy,
8. Infinite Universe.

From Atomic Nucleus to Metagalaxy

Man with his enquiring mind penetrates into the mysteries of the systems of a size both invisible to the eye due to their smallness and those which are extraordinarily enormous. It is interesting to compare how far he has penetrated in the two directions. Studying the systems of which he himself consists, man has reached the atomic nucleus, having a diameter of 10^{-13} centimetres, i.e. approximately 10^{15} times smaller than himself. Studying systems of which he himself is a part, he encounters a system 10^{15} times larger in the form of the solar system (strictly speaking, the now known diameter of our solar system is smaller, being only 10^{15} centimetres).

The diameter of the part of the Metagalaxy now known to us is about 10^{28} centimetres. In other words, we have penetrated into space 100 million times farther than into the microworld of tiny particles. Nevertheless, the properties of the greatest of the systems in the Universe have been understood by astronomers only on the basis of a study of the most minute particles known to physicists. But also in the study of this microworld the greatest assistance has been rendered by observation of processes in space, replacing those that cannot be made in the laboratory experimentally. The large and small merge in the unity of nature.

In fact, in order to clarify the structure and properties of matter it must be studied under all possible conditions. However, in laboratories on Earth we cannot for the time being create such varied pressures and temperatures as exist in stars and nebulae. Recall the study of the state of the interiors of stars – the white dwarfs, the discovery of helium on the Sun and its later discovery on Earth. Thus, winning over nature, man in a sense forces the celestial bodies to serve him. Their study makes it possible to understand the laws of nature more thoroughly.

Interaction of Galaxies

Both Herschels and Rosse (who discovered the spiral structure of a number of galaxies in his telescope) in the first half of the 19th century discovered and sketched nebulae connected by bridges or almost merging with one another. The nature of those nebular spots was unknown at the time, but in our time attention has been drawn to them by Zwicky. He described a number of unusual systems – galaxies, connected by narrow luminescent bands, which he called bridges or bars. It happens even more frequently that in closely spaced galaxies, or in one of such galaxies, there will be a bright tail extending from the main body. A similar tail, directed away from our Galaxy, was discovered in the Large Magellanic Cloud – a companion of our Galaxy.

The author, undertaking a special investigation, found several hundred systems where two or more closely spaced galaxies, connected by bridges, penetrate into one another and have tails submerged in a common luminescent fog or have a distorted spiral form. All this is the result of their interaction and is undoubtedly due to their common origin. It has been found that the bridges, tails and obscuring fogs, in which whole groups of galaxies are sometimes submerged, consist for the most part of stars and sometimes have an admixture of luminescent gas.

As an example, we will cite here a pair of elliptic galaxies NGC 750-1 connected by a thin bridge. The entire system was found to be submerged in an extensive stellar fog.

Another example is a pair of barred spiral galaxies 200,000 light-years across, which exceeds the dimensions of the galaxies themselves. One of the galaxies has a tail of almost the same length. In the case of the bright and relatively nearby galaxy M 51 in the constellation Canes Venatici, the bridge is one of the spiral arms of a larger galaxy. The author found a number of pairs similar to it and in 1959 and 1977 published two parts of an Atlas of Interacting Galaxies (more than 800 photographs).

One thing stands out, namely that tails are more frequent than bars and are normally brighter. This is

especially apparent in the system NGC 4676 found by the author.

The objects included in the Atlas began to be extensively studied in various countries. Of great interest is the Atlas of Peculiar Galaxies compiled by Arp in 1966. He used a 5-metre telescope to photograph half the objects from the Atlas of Interacting Galaxies, and some single peculiar galaxies. Owing to the five-fold scale and special measures, their striking features were brought out especially sharply. The author proved that most of interaction phenomena are not tidal and antitidal bulges, as was previously believed. Tide, on the contrary, must be stronger on a side facing a perturbing body. Besides, it is often noticeable with galactic pairs that they are dimmer on sides facing each other, and bright white stars are few there. Deformed are just the spiral arms whose origins still await their satisfactory explanation.

Examination of the 1959 and 1977 atlases of interacting galaxies assured the author that the process of fragmentation of galaxies still carries on. Some of the larger galaxies separate into two or three parts or small companions separate out at their edge. The initial phase of fragmentation is a "nest" of galaxies, which turns into their open groups. The ways in which galaxies interact are accounted for not only by the tidal action, but also by some unknown electromagnetic interactions. But bars and tails occur with gas-free elliptic galaxies as well. This strongly suggests that here we have some properties inherent to the system as a whole. These are some absolutely new properties, and between the galaxies there may be some forces of other nature than the well-known gravitation and magnetism.

There is nothing unusual in this possibility. In the world of molecules gravitation is dwarfed and molecular forces step upon the scene, and in the world of subatomic particles, in nuclei, we have nuclear forces and quantum processes. Beyond doubt, with the systems of ever growing size, instead of gravitation, which controls the motions of planets and binary stars, new forces or forms of interaction will appear somewhere.

Should these assumptions come true, it would appear that man would penetrate not only some special laws governing transmutations of elementary particles in atoms, but also the special laws of the largest among the known material systems.

The advent of fast computers enabled the motions of particles to be calculated that revolve in one plane about a galactic nucleus under the tidal perturbation produced by another galaxy flying by with a velocity close to parabolic one. It turned out that bars and tails must form, which however is insufficient to explain the variety of forms observed.

Radio Galaxies and Enigmatic Quasars

Normal spiral and irregular galaxies produce radio emission, comparable with that of our Galaxy. This radio emission and hot hydrogen content grow from Sa to Sc and irregular galaxies. This thermal radiation with a continuous spectrum is combined with the 21-centimetre radiation of neutral hydrogen and continuous nonthermal radiation caused by the braking of cosmic rays in the magnetic field of a galaxy.

Moreover, radio galaxies were found that produce a lot of nonthermal radio emission. The latter is due to the magnetic braking of a wealth of relativistic electrons and protons.

There is no proportional relation between the optical luminosity and the radio luminosity of a galaxy. The nearest radio galaxies are NGC 5128, or Centaurus A, and M 87 (NGC 4486), or Virgo A. Their apparent brightness is 8^m, their optical luminosity is large, and their distance is 30,000,000 light-years.

But the most powerful of the known radio galaxies and even the most powerful of extragalactic optical sources is the very distant galaxy Cygnus A. Its magnitude is 16^m, i.e. 1/1,500 the brightness of the next fainter one and is 20 times farther away than Centaurus A. The radio flux coming from it is 10 times that from Centaurus A, and the radio luminosity is about 4,000 times higher.

Each year now sees a discovery of fresh, ever more fainter radio sources, which are gradually identified with ever fainter, hence more distant, galaxies, and so the discovered radio galaxies grow in number quickly.

In the earliest days, it was thought that colossal radio emission is produced in collision of gaseous masses contained in the two colliding galaxies. But back in 1957 the author indicated that it is impossible for a number of reasons. It was gradually recognized that powerful radio emission is inherent in single galaxies and is not produced in collision of two

Radio galaxy NGC 5128 (Centaurus A).

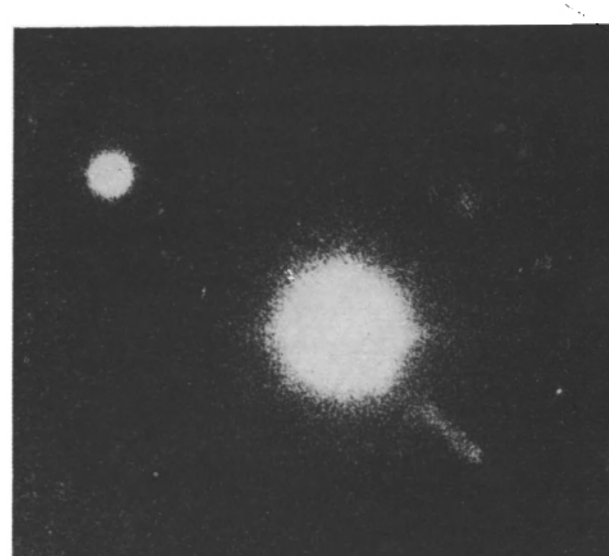
galaxies. On the other hand, it turned out that most extragalactic sources are binary. Radio components are on average separated by 600,000 light-years, an optical galaxy lying in between and sending out fainter radio waves. A radio emitting area is generally far larger than an optical galaxy. For example, the galaxy NGC 5128 is 90,000 × 70,000 light-years, nearly circular, whereas the associated radio source Centaurus A is much extended, its length being more than 1,500,000 light-years. And the two parts of Cygnus A are 200,000 light-years across each, the centre-to-centre distance being 300,000 light-years. The optical galaxy lying in between them is smaller by far. In 1966 four more symmetrical, nearly point radio sources were found near it, thus making its radio structure yet more complex.

The radio galaxies NGC 5128 and M 87 showed two departures in the optical range. They are both conventional elliptic galaxies, nearly spherical, but the former is crossed by an unusually thick and ragged dark bar, and the latter has a nodular appendix beginning at the centre, which was thought of as an ejection. Most radio galaxies produce bright, sometimes very wide spectral bands. Much effort went into attempts to find with radio galaxies some general features in configuration or spectrum. None have been found, however. We never know which of the galaxies will appear to be a radio galaxy. What is more, the author has shown that among the conventional, not radio-emitting galaxies, many appear to show the same features as the radio galaxies. Specifically, it was found that the so-called Seyfert radio galaxies also display very wide bright bands, suggestive of gaseous flows travelling at velocities as high as 5,000 kilometres per second. It turned out at a later date that NGC 1068 likewise is a radio galaxy, thus attesting to the similarity of appearance of conventional galaxies and radio galaxies. Galaxies are also known with strong, but narrow emission lines in their spectra, which are not radio galaxies though. They were discovered by the Mexican astronomer Haro.

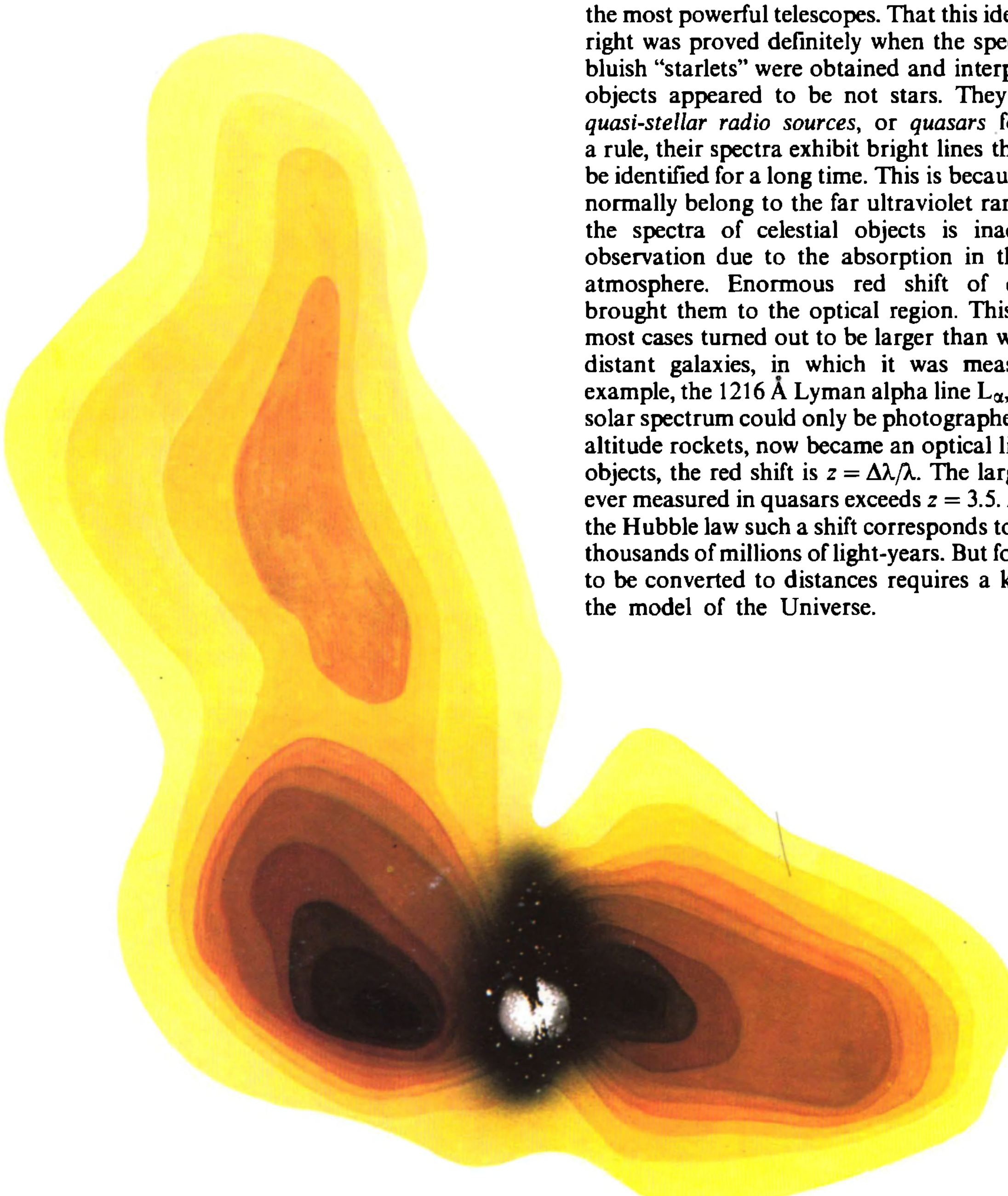
At present we already know many Seyfert radio galaxies. Their radio emission, just like powerful ejections and gaseous outflows from the galactic nucleus, seem to occur intermittently. It is these phenomena that are responsible for the wide bright bands in their spectra. The bluish colour of these galaxies is due to synchrotron, not stellar, glow of their small, but very bright nuclei. The same abnormally bright end of the spectrum is found in distant galaxies, of which Markaryan has discovered many. Some of them belong to the Seyfert type.

The most exciting development of the recent years has been the discovery by Sandage and Schmidt of the USA of unusual radio sources. After the positions of

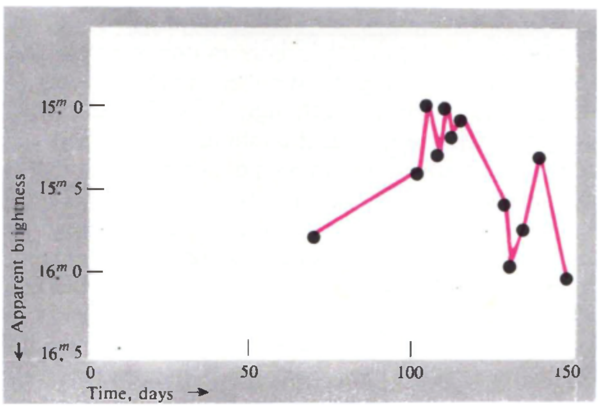
Quasar 3C 273.



The intensity distribution for the radio emission of NGC 5128 (Centaurus A).



the powerful radio sources were determined more accurately, some of them had to be identified with very faint point objects, indistinguishable from stars even in the most powerful telescopes. That this identification is right was proved definitely when the spectra of those bluish "starlets" were obtained and interpreted. These objects appeared to be not stars. They were called *quasi-stellar radio sources*, or *quasars* for short. As a rule, their spectra exhibit bright lines that could not be identified for a long time. This is because these lines normally belong to the far ultraviolet range, which in the spectra of celestial objects is inaccessible for observation due to the absorption in the terrestrial atmosphere. Enormous red shift of quasar lines brought them to the optical region. This red shift in most cases turned out to be larger than with the most distant galaxies, in which it was measurable. For example, the 1216 Å Lyman alpha line L_α, which in the solar spectrum could only be photographed from high-altitude rockets, now became an optical line. For such objects, the red shift is $z = \Delta\lambda/\lambda$. The largest red shift ever measured in quasars exceeds $z = 3.5$. According to the Hubble law such a shift corresponds to distances of thousands of millions of light-years. But for these shifts to be converted to distances requires a knowledge of the model of the Universe.



The variation of quasar 3C 454.3.

This is because in conceptual models of various types the red shift at large distances can vary nonlinearly with the distance. The same is true of the conversion of z into the velocity along the line of sight using the Doppler relation.

Quasars are mostly labelled by numbers from the Third Cambridge Catalogue of radio sources, abbreviated to 3C. The nearest and brightest quasar appears like a $12^m.7$ star, its red shift being 0.16 and velocity 48,000 kilometres per second. The number of quasars discovered grows amazingly rapidly. One of the most distant quasars, 3C 9 ($z = 2.012$), has a radial velocity of about 240,000 kilometres per second, a figure close to the velocity of light. Its distance was estimated at 9,000,000,000 light-years, in terms of years twice the age of the Earth. 3C 9 was one of the most distant objects in the Universe.

If the red shift for quasars has the same nature, this implies that their distances are enormous and their optical luminosity is 100 times that of the brightest galaxies and radio galaxies, their radio emission being similar to that of radio galaxies (10^{45}-10^{46} ergs per second), hence the name quasi-stellar radio galaxies or quasi-stellar radio sources. Like radio galaxies, their radiation must be of synchrotron nature, i. e. must be caused by bremsstrahlung of relativistic electrons.

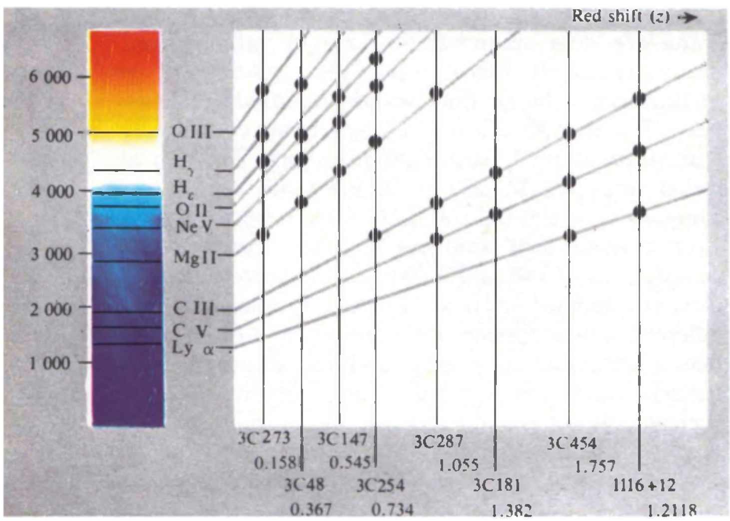
The rapidly growing accuracy of measurements of angular sizes of radio sources enabled astronomers to find that many quasars have radio-emitting areas with diameters of fractional seconds of arc, and often less than $0''.1$. Optically, too, they are indistinguishable from stars. It is impossible to measure optical diameters less than $0''.5$ due to atmospheric perturbations. It follows that the optical size of quasars is not larger than several hundreds of light-years. But near 3C 273 and 3C 48 there are exceedingly faint straps about 200,000 light-years long. The optical stellar image of 3C 273 is, in the radio range, surrounded by a faint halo, and $19''.5$ away from it a faint strap is visible that produces 10 times stronger radio emission. In 3C 48, a 16^m "starlet" is surrounded by five nebulae lying as far as $12''$ away. Quasars are thus different in their appearance, but no one of them looks like a conventional galaxy, and the dimensions of quasars are extremely small. It may well be that the optical radiation of quasars is dominated by synchrotron radiation.

A new exciting development was the discovery that in quasars both apparent brightness and radio emission change significantly. The profiles of the bright lines of hot gases appeared to change as well.

Another break-through came in 1965: Sandage discovered in the area of the Galaxy's pole many very faint blue starlike objects, that resemble quasars in colour. He obtained spectrograms of six of them. One spectrum belonged to a normal, relatively nearby star, two spectra contained no lines, and in the other three cases bright lines were found with large red shifts, just like in quasars, although no radio emission from them has yet been found.

Sandage called them *quasi-stellar galaxies*, or *quasags*. From his counts of the blue objects he concluded that they must outnumber quasars by hundreds of times. Further studies have shown that most blue objects at the Galaxy's pole are the bluish stars of different types belonging to the outskirts of the Galaxy, but quasags are really much fewer, although still more than quasars in a unit volume. Zwicky believes that Sandage's quasags are identical with those of his extremely compact galaxies, which are bluish and show bright spectral lines. (We here speak about identical types, not individual objects.)

Quasars are assumed to be a short-term phase of the fast development of quasags, which is why powerful radio emission is only observed in some of them, when we identify them as quasars. In any event, the discovery of quasars and quasags was the most fascinating development in astronomy of modern times. These are some absolutely new kinds of celestial



Positions of emission lines of quasars with various redshifts.

objects possessing some mysterious properties, which perhaps lead us to a discovery of the most general laws of nature.

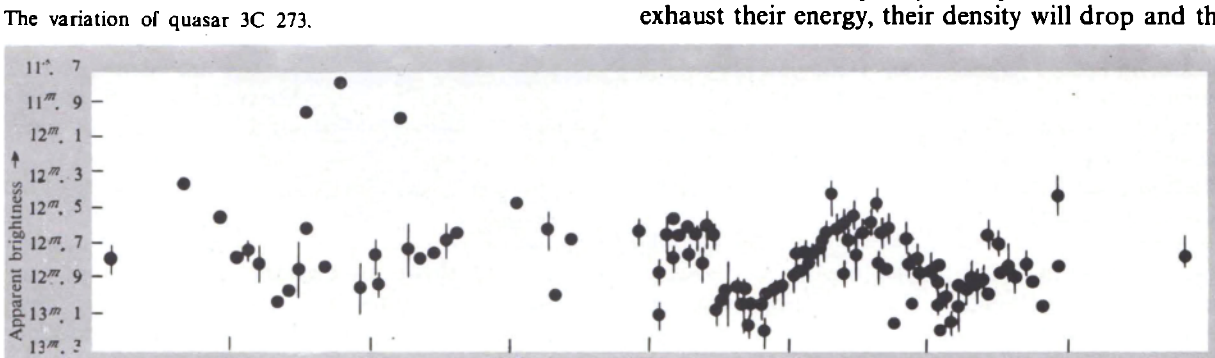
Explosions in Island Universes

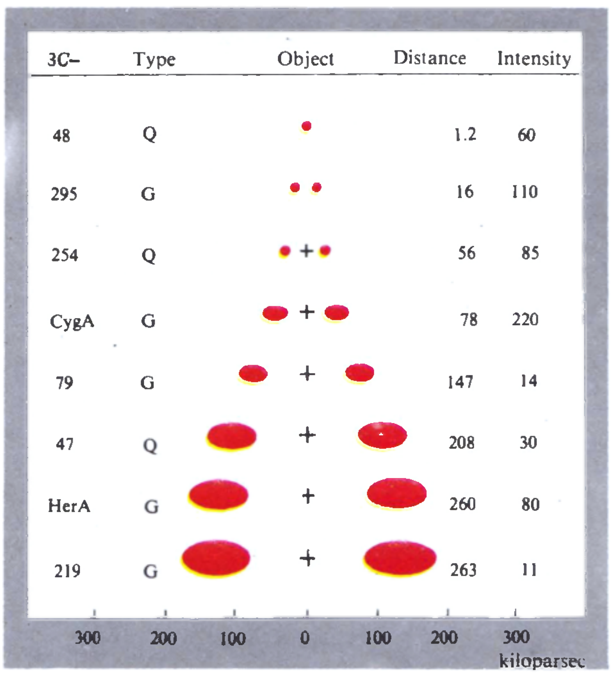
Getting acquainted with the latest discoveries in astronomy you cān stop surprising at anything. Take, for example, the discovery of giant explosions in the solar atmosphere. But these, too, pale before the explosions on UV Ceti flare stars. Also worth mention are the nova and supernova explosions. And here we turn to the story of explosions in island universes.

A large and beautiful spiral galaxy in Ursa Major, M 81, has a companion. This is an uncomely, longish nebula M 82, having, as it were, "rugged" edges. It attracted no attention, although differed from normal irregular galaxies in that it contained much dust and, at the same time, contained no hot blue giants, although its spectrum was of A-type. It was M 82 that was a prototype of irregular galaxies Ir II. But in the last two decades it became the most "fashionable" galaxy. It was made famous by Sandage and Linds in 1963.

The pictures of M 82 taken in the red hydrogen line H_α distinctly show long hydrogen filaments, spreading on either side of the centre. They are oriented normally to the galactic plane, which forms a small angle with the line of sight and therefore appears longish. It turned out that the gas in these filaments moves the speedier the farther away from the centre it is. Apart from the red hydrogen filaments, the pictures show bluish filaments yielding a continuous spectrum in polarized light. It is quite obvious that this is due to the flows of fast electrons producing synchrotron radiation both in optical and radio ranges. Also they ionize hydrogen atoms in collisions. Gaseous flows are directed to the poles of this rotating galaxy and not in its plane, which is caused by the drag produced by stationary gases that have been there before. The gas there is mixed with light-absorbing dust visible to us. The gases in M 82 fly apart with a kinetic energy of about $2 \cdot 10^{55}$ ergs, and during the one and a half million years elapsed it emitted as much as 10^{56} ergs. This is a million times more than the energy released in a supernova explosion. The ejected gas propagates through 10,000 light-years from the centre and beyond the confines of the galaxy. The gas and electrons will exhaust their energy, their density will drop and they

The variation of quasar 3C 273.





Relative sizes, distances and intensities of components of extragalactic radio objects.

will scatter. The explosion and the radio emission produced in it are fairly short-term events as compared with the galaxy's age, which is estimated at 10,000,000,000 years.

During the 15 years after Sandage and Linds's studies of M 82 it was believed that the phenomena observed were produced by a central explosion in it, although it was still unclear what actually exploded. M 82 was referred to as an "exploding galaxy" and hypotheses were in vogue that drew a parallel with atomic explosions. The concept of "star formation waves" was in current use. Some authors, it seems, did not take the trouble of more specifically defining the hypothesis and suggested to treat M 82 not as an exploding galaxy, but as one in a state of active formation of stars and hot gases as a result of its interaction with the spiral galaxy M 81.

Even before the examination of M 82 it was supposed that binary radio sources, between whose components lies an optical galaxy, are formed by explosions. A galactic explosion ejects two huge dust clouds filled with relativistic electrons as a sponge with water. From momentum conservation law, the clouds move in the opposite directions, and the old gas in the galactic plane makes them divert to the rotation poles. After the clouds emitting the synchrotron radio radiation leave the confines of the galaxy, we see two extended radio sources on either side of the galaxy. In radio galaxies the energy release is yet more impressive than in M 82. During the time Cygnus A has been in the stage of radio galaxy, which is estimated at a million years, it has radiated $3 \cdot 10^{58}$ ergs. The energy of synchrotron radiation together with the kinetic energy and the energy losses give the explosion energy in Cygnus A to be 10^{60}-10^{61} ergs. It was equal to the energy of conversion of 1,000,000,000 solar masses of hydrogen into helium. The colossal amount of energy liberated nearly instantaneously and the unknown physical mechanism of its source are the main puzzles of the origins of both radio galaxies and quasars, whose energy is similar.

The author noted in 1956 the similarity of spectra of the Seyfert galaxies and some of the radio galaxies. Now this similarity attracted more attention. It turned out that the violent release of hot gases from nuclei of the Seyfert galaxies is an explosion. Their nuclei are starlike, i.e. are small. Further, at the centres of some of the Seyfert galaxies point radio sources were found. It is therefore said that at their centres we find sorts of small quasars. Quasars are, as it were, powerful exploded nuclei of the Seyfert galaxies, but without the entourage of a stellar galaxy.

It is especially difficult to account for quasars. In addition to the difficulties of finding suitable sources of energy and mechanisms of liberation and conversion of this energy into the energy of relativistic electrons and their total motion, we have here the difficulty of explaining their small dimensions. The fact is that they cannot be stellar systems. A large ensemble of stars cannot undergo the observed fast oscillations of the total brightness and radio emission. This must be some single huge body. Initially it was suggested that in a dust cloud of 10^8 solar masses there occurs under the action of gravity a catastrophic compression, a so-called collapse, and a superstar forms. The compression releases a colossal amount of gravitational energy. But it is as yet unknown how it can convert into the energy of relativistic electrons. According to this hypothesis quasars were quickly called superstars. The hypothesis, however, did not

M 82, or NGC 3034, an irregular stellar system known as radio source 3C 231.

M 82 (H_α-picture).

gain much recognition, and now a dozen various hypotheses have been put forward to account for quasars, and these are now vividly discussed. Among them, there is a group of hypotheses that attempt to treat quasars as nearby objects, and their red shift is explained otherwise than by the effect of their being distant. These attempts can hardly be a success. We cannot here dwell too much on the various quasar hypotheses, of which none was recognized. Hopefully, the mounting factual evidence will eventually lead to a correct explanation.

Note that most scientists believe that stars and galaxies emerge as a result of condensation of tenuous gas. Speaking about galactic explosions, they normally avoid discussing the issue of what explodes in the final analysis.

A more consistent point of view is that of Ambartsumyan who assumes that both stars and gas emerge in explosions of some superdense matter. He believes that the nuclei of some galaxies include within a small volume a prodigious mass of superdense matter capable of dividing in an explosion-like manner and give rise to pairs and groups of receding galaxies. Minor ejections produce companion galaxies. Radio galaxies, and maybe quasars too, are thought of as galaxies whose nuclei are in the process of catastrophically dividing. As mentioned, much evidence is available to support the assumption that many groups and even clusters of galaxies disintegrate, although it is unknown where the energy required comes from. But the same question remains in force as regards the radio galaxies and quasars discovered later. True, so far no large and superdense masses have been found in galactic nuclei, but now the possibility of such a discovery seems less unlikely than it appeared earlier. Ambartsumyan was the first to attract attention to the activity of galactic nuclei and their great evolutionary role. At an early date, Ambartsumyan paid attention to the ejection from the central part of the radio galaxy M 87 (similar ejections

are observed in other galaxies). The emission from this ejection turned out to be synchrotron emission in the optical range and connected with the radio emission. The nuclei of galaxies, their radio emission and other phenomena are studied at the Byurakan Observatory.

Is There a Boundary of the World and What Is Beyond?

Long before the enormous distances to galaxies have been established, humanity kept asking: is there a boundary of the world, and if so, what is beyond it? The study of the world as a whole is the subject of cosmology. This subject also concerns philosophy, mathematics, which deals with the notion of infinity, and astronomy, which examines specific celestial bodies. The question is exceedingly complex. So the philosophy of dialectical materialism teaches that matter and motion are eternal, although they undergo conversions. There is hardly a scientist who doubts the infinite variety of natural phenomena, which are invariably material in nature, although the proponents of idealism attempt to interpret any new phenomenon that still awaits its explanation in an idealistic way. But they are losing ever more ground with each advance of science. Now, it seems, only few scientists really believe that the Universe has a boundary, a "wall". However, the question of whether or not the Universe is finite and what the properties of space are can be answered by observations in space.

In schools we study the Euclidean space in which two parallel straight lines never cross each other. But the great mathematician Lobachevsky proved that a space is conceivable with other behaviour. Later Einstein showed in his theory of relativity that a real physical, not abstract, space filled with matter can have a measure of curvature caused by the existence of matter. The Soviet scientist Fridman, and later other scientists, developed mathematical model of universes relying on the theory of relativity. There are many such models and most of them deal with an endless, but finite Universe. A combination of endlessness and finiteness is generally exemplified by a ball. Having a finite volume, it has no boundaries for a creature travelling over its surface. The ball's dimensions can grow, decrease or pulsate, still remaining finite.

Theoretically, the behaviour of a finite Universe is dependent on its average density and the homogeneity of the density over the volume. Turning to observations, we can as yet study the part of the Metagalaxy that is often and without ground identified with the Universe as a whole.

We have found, by examining their red shifts, that galaxies recede from one another and they do so the quicker, the farther away they are from one another. We have some information about the masses of galaxies and their distribution in space. The Metagalaxy seems to expand, but which of the Universe's models describes it the best? It appears that this can be found out, by establishing the relationship of the amount of red shift to the distance of a galaxy, if the latter is determined in an independent fashion (and not from the same red shift). Also this purpose may be served by the distribution of extremely distant galaxies (or radio sources) in space. As it was said above, the distance of galactic clusters can be determined from the apparent brightness of the brightest galaxies in them. The results of the observations are compared with the theoretical predictions for different models. Our current understanding of the workings of the Metagalaxy and the accuracy of the data available are as yet insufficient for a definitive conclusion to be made. Still most of the authorities are now led to conclude that the Metagalaxy is finite and expands with a deceleration that produces mutual pull. Probably, pulsations do occur, if not of the Universe as a whole then of the Metagalaxy, and at some time the expansion will be replaced by contraction.

The expansion of the Metagalaxy suggests that several thousands of millions of years ago its volume was so small that galaxies could not exist as individual objects. This by no means implies that it was at that time that the world was created, as idealists used to maintain. It was simply that matter then existed in another form. Matter can convert in an infinite variety of ways, and it not always has been and not always will be in those forms in which we observe it around us now.

Is Any Communication Possible with Extraterrestrial Civilizations?

"The news with signals from Mars! Marconi's receivers receive telegrams from Mars!" So wrote newspapers at the turn of the century. At the time much fuss was in the press about the observations of

Martian canals and their supposed artificial origin. This sensation set the scene for the reports that in South America an engineer of Marconi's firm heard some crackling on radio. But these appeared to be atmospherics, now known to each radio enthusiast, which are produced by electric discharges in the terrestrial atmosphere. Later on no radio signals of artificial extraterrestrial origin were found. The famous French popularizer of science Flammarion joked that the Martians should long have lost desire to attempt to establish radio communication with the Earth. In fact, as early as millions of years back the supposed Martians could have sent out signals, which neither brontosaurs nor pterodactyls answered. No more sociable were the people of the Stone Age, the Egyptians of Ramses's time who erected the pyramids, and others.

And several years ago some science-fiction writers produced a sensation that certain of the religious stories, cave pictures and amazing structures of the ancient times testify to... the Earth being visited by "guests from space", dwellers of other worlds.

In principle, the Earth could have been visited by extraterrestrial astronauts. But what is claimed to be the traces of these visits border on anecdote. So the story of the ascension of the saints to the Heaven allegedly accounts of the visits of extraterrestrial guests. And hitherto unexplained stupendous stone structures prove nothing whatsoever. How could possibly the ancient people on the Easter Island erect the giant idols? That has long been a puzzle. But the descendants of those people showed to Thor Heyerdahl how this can be done without any machinery, let alone astronauts. This invention, just like the tale that the Tunguska object was no comet or meteorite but an atomic spacecraft from Venus or Mars, bursts like a soap bubble even because science has established that no civilizations are possible on any of the planets of the solar system. A civilization that would be able to sent their interstellar spaceships may only be on those planets that revolve around distant stars. But even the nearest star is four light-years away and, as we shall see later, the nearest technical civilization can only be expected at hundreds of light-years. A flight to visit us in a spaceship, even in a nuclear-powered one, is out of the question. Even the so-called photon rockets, which travel at near-light velocity, would take hundreds of years for their round trip. True, according to Einstein's theory of relativity, time (as understood by mechanics) is slower with accelerated motion. It has been calculated that if a ship travels first with a uniform acceleration of 10 metres per squared second and then it decelerates following the same law, the clock on the ship would be so slow that it would indicate the passage of only several years taken by the ship to cover the distance to a distant star.

But it is quite unlikely that such ships will ever be constructed. What is more, a collision of the ship travelling at such a velocity with interstellar medium will be disastrous for it. Although these estimates seem to be valid for hardware, they are not suitable for living and intelligent beings.

Well and good, suppose that such spacecraft could be constructed, all the difficulties and hardships of the flight could be surmounted and the astronauts could survive. But during the journey hundreds or thousands of years would pass on Earth. The society on Earth would have developed so that the information brought back by the astronauts to Earth would be as valuable for the then people as the description by Jenghis Khan of what he had seen when flying by other stars for us.

But this is not the whole story. Recall how many launches were required to land safely the first automatic station on the Moon. Yet more effort went into the mooning of the astronauts. Each fresh start of a spacecraft with an interplanetary station allows us its drawbacks to be found and taken into account in further launchings. With interstellar travel, however, we would wait for hundreds of years for the results. Would such ventures be reasonable? With the colossal energy and other expenses to boot.

Even rejecting the idea of visiting other civilizations, could we establish some sort of contact with them by exchanging radio signals? We could evaluate the chances by making simple calculations.

To begin with, what distances can on average be expected between civilizations that are able and willing to enter into a radio communication with one another?

The Galaxy has about $N = 10^{11}$ stars, a fairly hard fact. But what follows is sheer conjecture. The number of communicative civilizations in it, n, will be

$$n = N \times P \times E \times B \times I \times T,$$

where P is the share of stars having planets, E is the share of the planets suitable for life, B is the probability

that under suitable conditions life does emerge, I is the probability of the life producing a sufficiently advanced civilization, and T is the lifetime of this civilization. After all, we are interested in those civilizations that are contemporary with ours, not past or future.

We do not know the percentage of stars having planets. As we have already discussed, with some of nearby stars invisible satellites of small mass were found, and the effect observed is actually produced not by a single sattelite, but by a group of bodies of smaller mass, which is about the mass of the larger planets of Jupiter type. In consequence, that planetary systems exist, apart from our system, is beyond doubt, but the question of whether all stars have them and which have and which have not still awaits its time. As an optimistic estimate, we will consider that all the stars have planets, i.e. $P = 1$.

What percentage E of the planets is suitable for life? It is anyone's guess, considering that we only know the planets of our solar system and insufficiently at that. Different authorities gave widely varying estimates: from one millionth to one twentieth. Accordingly, in the Galaxy the number of planets suitable for life is estimated to be from 10^5 to 10^{10}. Suppose also that, given these conditions, life on such planets is bound to emerge, i.e. we put $B = 1$. Again this is an overestimation.

And on what percentage I of inhabited planets the development of life results in the emergence of intelligent creatures, who besides are at such a stage of technological development that radio communication with other worlds is well within their capabilities? It is also unknown what percentage of such civilizations will spend enormous amounts of energy for a very long period of time, actually for hundreds of years, waiting for an answer from other planets. In all probability the answer will not come because either the addressee will not understand the question or to send it out with such a delay will be pointless. It should be taken into account that the difficulty may be due to the fact that there may be no common language in the direct sense of the word, just like between two peoples on Earth, and in the sense that the difference in the levels of the two civilizations may be too large. After all, there may be civilizations whose level of knowledge, concepts and psychology may differ even more from ours than ours differ from those of ants. Only under the most favourable conditions can we assume that $I = 1$ and that we need not introduce other fractional factors into the formula.

Lastly, also uncertain is the estimate of T. As mentioned above, we are only interested in contemporary civilizations. Life on Earth has been developing for about half a million years, but the idea of communication with other civilizations occurred to us only several years ago. It would be of no use to appeal to mankind from space for the half a million years. Should we tomorrow receive unexpectedly some artificial signals, we would not be able to answer anything sensible. Our today's civilization (with its radio engineering and aeronautics) is rather young. How long will it last? In recent decades a new danger of destruction of humankind has appeared, caused by the fact that science and technology have made so great strides that the mind cannot grasp it yet. This is the danger of self-destruction. But will all the civilized worlds end through self-destruction? There is no way of answering the question. (Note that as a result of interplanetary visits lethal bacteria or viruses may be brought in, which may destroy the civilization before some means will be found to combat them.) If all civilizations self-destruct, then $n = 0$, and that is the end of the story. The lifetime of a civilization is estimated to be 10^2-10^9 years, and so the number of contemporary civilizations in the Galaxy will range from 10^2 to 10^9. Therefore, even for upper estimates giving exaggerated results, the latter remain exceedingly uncertain, being different by 10^7 times. The above agrument illustrates a scientific approach to the problem.

Proceeding from a well-known average distance between stars (seven light-years), we can calculate the average distance D between civilizations depending on the assumed number of contemporary civilizations d. We thus obtain:

d	D, light-years	d	D, light-years
10^2	7,000	10^6	320
10^4	1,500	10^9	32

The lower limit of the average separation between civilizations is thus hundreds or thousands of light-years. How long will a light or radio signal take to cover the distance?

At such distances we can only hope to achieve a communication using highly directional transmission. Direct visits, as said above, are unlikely. Also meaningless is communication in those cases where (as in the first line of the table) the duration of information exchange is longer than the life of, at least, one civilization. In general, the probability or meaning of the two-way communication is small. There thus remains one-way communication: transmission without reception or just reception. At present, mankind can only think about reception, not transmission.

Can present-day radio telescopes receive signals from nearby civilizations? And how? How then can we decipher the signals?

When radio signals are sent out in a narrow beam using modern facilities, they can be detected if the transmitter has a power of 100 kilowatts, a frequency band of 10 kilohertzes and an antenna 100 metres in diameter, when the transmitter is separated from us by a distance of the nearest stars. But at what frequency are we to expect the signals? It is highly unlikely that the transmission will be simultaneous at all the wavelengths, otherwise it would be extremely difficult to distinguish it from the natural radio emission of celestial bodies. It was supposed that a civilization would choose a wavelength of 21 centimetres, since it is a line of cool hydrogen, which is extremely abundant in the Universe and must be studied by all civilizations that are advanced enough to come into radio contact with other worlds.

In 1960 in the USA a 27-metre radio telescope was used to "listen" to cosmos up to a distance of 16 light-years. Within this range there are three stars near which some inhabited planets could be expected. Several months of "listening" to these stars gave no result. But the telescope was insufficiently sensitive – chances of success would be much higher if the range were increased up to hundreds of light-years. On the other hand, near 21-cm line there is too much cosmic radio emission, which could suppress artificial signals. Civilization might as well select another wavelength, including one that is not passed by the terrestrial atmosphere, but a radio telescope cannot simultaneously receive signals within a wide range, another complication. It is also impossible to observe each of 10,000,000 stars lying within 1,000 light-years, since it is impossible to establish beforehand near which of them there is a beaming civilization.

For signals to be able to reach us from 1,000 light-years a transmitter is required with power of the order of millions or thousands of millions of kilowatts. According to estimates, we could construct it on Earth for about $ 200,000,000,000.

It was established that the communication range is also dependent on the signal duration and grows with it, whereas increasing the frequency range in transmission (hence making the reception more difficult) also increases the signal transmission speed. It may well be that while the experience and knowledge of one civilization were received and deciphered by another civilization, the former would cease to exist. But the data would help the recipient civilization.

What language can be used for the communication and how it is to be deciphered – there is no knowing, but the question is being discussed. What is not discussed, however, is the question of the compatibility of the levels of civilizations and the possible difference in their psychologies. It is well known that even on Earth peoples of different countries and even scientists working in the same field not always achieve adequate understanding. Well, and if the levels of knowledge and psychology are as different as those of man and monkeys, or birds and fish? Could they understand each other, then?

In order that we may be sure that the signals received are of artificial origin, they must show a measure of regularity, they must be intermittent, polarized and clearly come from a very small volume.

It was recently assumed that there exist supercivilizations that consume 10^{15} times more energy than the terrestrial civilization, i.e. about 10^{33} ergs per second. For this to be the case, they would have to use completely the energy coming from their stars, for example, by surrounding them with an artificial sphere to retain all the energy. It was shown that such a sphere would be unstable and hardly possible in principle. Such a civilization could

beam signals within a wide frequency range (10^6-10^9 hertzes) and even to the most distant stellar systems. Now they even toy with the idea of supercivilizations with an energy consumption of 10^{45} ergs per second, which have mastered the energy of a whole galaxy.

When in 1965 it was suspected that one radio source is periodic, the author of the above hypothesis was very quick to inform the world that signals of another civilization had been detected on Earth. But after a further study the object appeared to be a quasar, i.e. one of quasistellar galaxies characterized by an extremely powerful radio emission and separated from us by thousands of millions of light-years. From such a distance, only a fantastic supercivilization could have beamed signals. We already know that many of such objects change the intensity of their radio emission and light for natural reasons. Although it is hard to believe that such a civilization can exist, attempts to detect adequate signals from a more nearby civilization should nevertheless be continued – just in case.

6. Birth, Life, and Death of Stars

How Old Are the Stars and the Milky Way?

We are not satisfied, of course, with those legends which tell of the year of creation of the world. According to the legend of Judaic priests, the world was created 5,765 years ago (considering that this is now 1985). However, the Byzantine church reckoning of the years, which was adhered to in Russia until the time of Peter I, considers that 1985 is 7,491 years from the creation of the world. Who is to be believed? We don't have to believe either one, because the petrified remains of plants and even animals, many millions of years old, have been found in rocks.

Dispensing with religious legends, we will attempt to reestablish the history of our planet by scientific means, and also the history of other celestial bodies, although no one has established definitely the date when they appeared in the Universe. It can be said that the birth of these worlds was no matter of hours or years, but of millions of years and the "exact" age of bodies in the Universe is a general concept, without literal meaning.

As we already have mentioned, the decay products of radioactive elements contained in rocks can to a certain degree be used to measure their age. For example, it has been established that the age of the most ancient rocks in the Earth's crust is about 3,000,000,000 years.

In dealing with such large periods it is possible to use as a unit time not the period of revolution of the Earth around the Sun, but the period of revolution of the solar system around the centre of our stellar system – the Galaxy. This period is about 250 million Earth-years. If this unit of time is called the "cosmic year", the age of the Earth's crust will be about 20 cosmic years.

How then can we determine the age of the stars and the Sun, which we are accustomed to think of as the "father" of the planetary family and the oldest in age? The method based on radioactive elements, which has been applied successfully on Earth, cannot be used for either the Sun or stars. We cannot obtain samples of their matter, and even if we could, it would tell us nothing because these bodies consist of hot and continuously mixing gases. It is clear that it is very difficult to estimate the age of the Sun and stars with assurance, but the problem is not hopeless. Various methods for making such an estimate have been devised recently, which agree rather well with one another, and when the estimates of individual observers agree with one another on the basis of different methods the ages must be regarded as correct.

The ages of stars and our stellar system itself are revealed by the structure of the stellar system and study of stellar motions.

The fact is that our stellar system has a definite and rather complex structure and is not a chaotic mass of stars. It is not a crowd of disorderly individuals, but is like the order of battle of an army unit, with a complex structure and subordination. Just like fallen soldiers are replaced in the ranks by reserves, so also the stars ejected from any cloud in the Milky Way are replaced by others, so that the general picture does not change in the course of tens of cosmic years. We refer to such a situation as the "dynamic equilibrium of a system". This equilibrium can be disrupted.

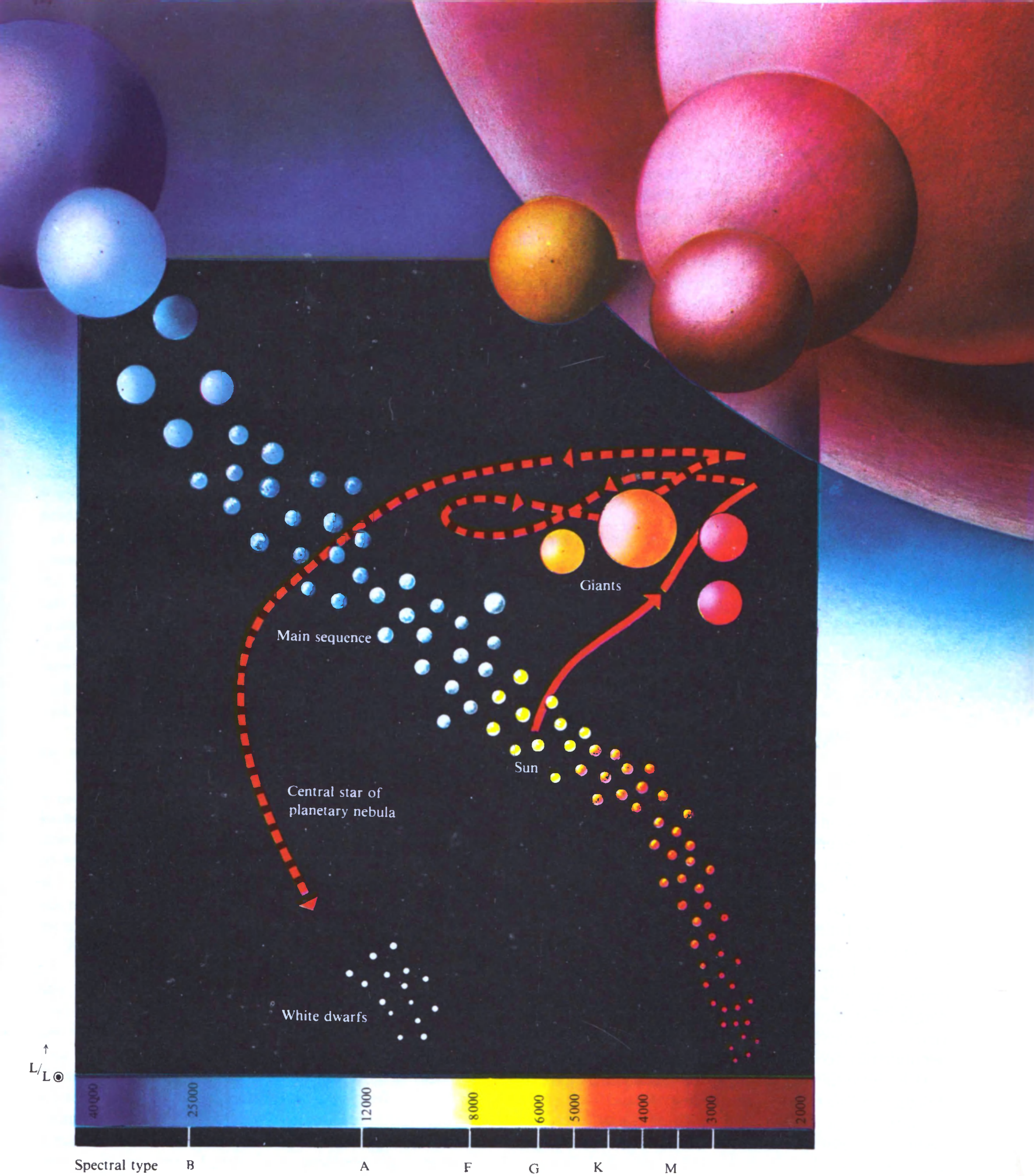
The distances between stars are so great in comparison with their dimensions that we can obtain a model of the stellar system if in place of stars we conceive of several dust specks in the great auditorium of the Bolshoi Theatre in Moscow. What are the chances that they will collide?

Knowing the distances between stars, their velocities and dimensions, we work out that a collision between the Sun and any star may happen once in 200,000,000,000,000,000 years or once in a milliard cosmic years.

In our stellar system, which has up to 200,000,000,000 stars, a collision between any two stars occurs on average only once in a million years. Possibly it would be simpler to say that collisions between stars, which, of course, would lead to the destruction of their planetary systems, for all practical purposes never occur.

Even close encounters of stars, during which their paths would be changed significantly under the influence of mutual attraction, occur with only the greatest rarity. Distant passages of stars at distances of around one light-year (i.e. less than the mean

A possible development path (final stages) for a star of one solar mass in the Hertzsprung-Russel diagram.


Giants
Main sequence
Sun
Central star of planetary nebula
White dwarfs
L/L⊙
40000
25000
12000
8000
6000
5000
4000
3000
2000
Spectral type
B
A
F
G
K
M

distance around stars, which, as we already have mentioned, is about 7 parsecs) occur quite frequently, changing the direction of their paths by an angle of one minute of arc. Such encounters produce a slow but inevitable disruption of that order in the stellar system which was established at the time of its "birth". As they say, "drops of rain wear away the stones". We have the same type of effect here as well. In the course of one cosmic year the total effect from stars passing by the Sun at a distance of ten to fifteen light-years will be the same as a single encounter with another star at a distance only 3 or 4 times greater than the radius of the orbit of Pluto (40 AU).

The motions of stars in the Galaxy reveal curious peculiarities associated with their physical characteristics. Massive stars with a corresponding "solidity" move in one direction or another, at the same time revolving around the centre of the Galaxy. However, red dwarfs, which are cool stars with a lesser mass, rapidly move about in the Galaxy in all directions, disrupting order but also participating in the general rotation.

Encounters and the mutual attraction of stars tend to redistribute the energy of motion uniformly so that the energy of each of them (equal to the product of its mass and the square of velocity) will be identical and all massive stars will move more slowly than light stars. In a chaotic mass of gas molecules the same thing occurs, as is well known to physicists, but in this case the molecules exchange velocities by direct collisions. Such a uniform distribution of energy between stars (or molecules, it makes no difference) can set in only after the passage of a sufficiently long period of time.

It is obvious that the Galaxy either has rotated an insufficient time for such an energy redistribution to set in or the stars are of different ages and certain of them are forming even now. If the Galaxy had existed in the form in which we now see it and had rotated for more than tens of thousands of cosmic years, various stars would not reveal their individual peculiarities to such a high degree.

It is obvious that acting over the cosmic ages, the passages of stars by one another should scatter star clusters. Among such star clusters there are rather open ones, not particularly dense, such as Pleiades and Hyades. Hyades include up to 150 stars, scattered to a distance up to 15 light-years from the centre of the cluster, which is 130 light-years away.

An analysis made by Ambartsumyan reveals that the Hyades cluster has been preserved from scattering for at least tens of cosmic years. Later the centre of the Galaxy will begin to threaten its integrity considerably and in approximately 50 cosmic years the Hyades group will cease to exist – its members will be scattered in the Galaxy.

But the Hyades is already an extremely open cluster. Others of its companions – Pleiades and Praesepe – have a mean density ten times greater. Their stability is higher, but they will not be able to withstand the attraction of the centre of the Galaxy longer than a hundred cosmic years, and after this period of time not a single open star cluster should remain in the Galaxy. In our time the process of their destruction is proceeding so rapidly that such clusters having developed more than 50 cosmic years ago no longer exist. That means that these characteristic members of our stellar system do not exist longer than 50 cosmic years, but the age of the Galaxy itself and the other component stars, including the Sun, can be greater. The Sun, many stars and our stellar system as a whole can be just a few times older than the Earth's crust, which has an age of 15 cosmic years. At the same time, hot giant stars should be young, not older than 1 or 10 million years, and certain of them are being formed at the present time.

The above sources of stellar energy can maintain their radiation over such a period of time. We still do not know of an energy source that could maintain their radiation over a period many times as great, but a shorter solar lifetime is unlikely on the evidence of the Sun's effects we find on Earth.

We see that the age of celestial bodies is venerable and the history of our science is just a brief blink in the lifetime of a star, so that we cannot note any changes of stars. During this time not even the slightest wrinkle has developed on their brows. This is a great hindrance to us when we seek to find how stars are born, how they develop and age. The fact that the bright Sun cannot tell us about its "dark" past (dark for us) makes it impossible to predict its future with assurance. A knowledge of the sources of stellar energy is of primary significance in such attempts at prediction of the future of our Sun.

These sources were discovered in the thirties, and we will now discuss them; only recently the poets could say nothing on this subject but:

In the depths of the unknown,
Full of mysterious forces,
Move millions
Of age-old bodies...

(I. Nikitin)

Where Does Stellar Energy Come From?

They say "we eat to live". The assimilation of food gives to living beings the energy they expend in motion. Any machine used for work requires "feeding" in the course of its work. Lathes require electric current and electric power stations require coal – the petrified plants of the distant past. These plants required solar heat and light, but what does the Sun itself require? How do stars expend such colossal quantities of energy? It must be made up for, because there is no perpetual motion in nature and unfortunately certain ill-advised inventors still do not know this.

If the Sun consisted of the best coal and this was burned in an adequate quantity of oxygen, it would burn up entirely in 1,500 years, but, of course, no oxygen would be available for this process.

It once was thought that solar energy is made up by the falling of meteorites on its surface. Their kinetic energy would be transformed into heat in falling, thus maintaining solar radiation. An unbelievable number of meteorites should fall on the Sun and they should increase the solar mass so rapidly that this would be noticeable. This method of supplying energy would help the Sun no more than it would help us to boil a barrel of water by setting hot flat-irons on its lid.

The Sun should receive energy from within , as we now know from all data concerning the nature of the Sun.

A series of hypotheses have been proposed and later refuted. We will list them briefly, primarily so that no one will undertake the vain effort of repeating the errors of the past.

Solar energy could be supplemented by the compression. In the process, the energy of attraction towards the centre would convert into thermal energy. It has been calculated, however, that even if the Sun was once infinitely large, its compression to its present size would suffice to maintain its energy for only 20 million years. But it has been demonstrated that the Earth's crust has existed and has been illuminated by the Sun for a far longer period of time. Compression can take place and probably has taken place, but it does not serve as the primary source of solar energy. Is it possible that the interiors of stars consist of radioactive elements such as thorium, uranium, and radium? Their decay would produce heat. If the Sun consisted entirely of radium (by the way, only a few grammes of it have been produced on Earth), it would radiate far more energy. But despite initial wastefulness, inevitable in radioactive decay, the intensity of its radiation would fall off too rapidly. Radium could not maintain the solar radiation as long as necessary. Modern physics and the theory of the internal structure of stars do not permit the existence of heavy "superradioactive" elements (unknown on Earth), concentrated in the interior of the Sun.

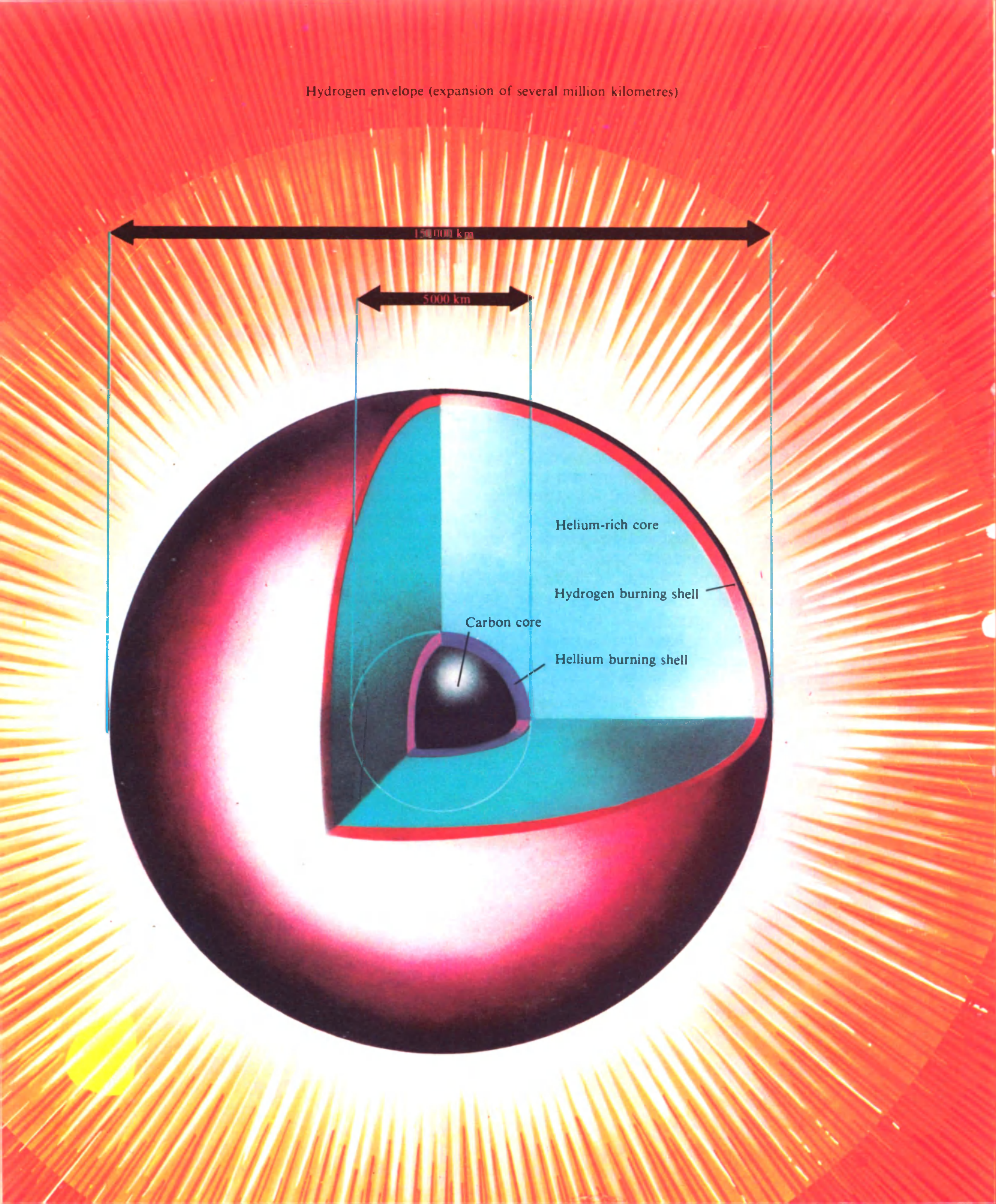
Fortunately, the nuclear physics, which originated in the 1920s, has revealed to us the source of stellar energy, which agrees well with astrophysical evidence and in particular with the conclusion that stars mostly consist of hydrogen.

Have you heard that hydrogen burns? Hydrogen in stars "burns" and gives them the necessary energy, but this is not the combustion that is actually the combining with oxygen. Combustion is a chemical process, but the energy of chemical reactions is inadequate for maintaining solar heat.

On the other hand, considering the extraordinary heat within the stars, no molecules or atoms can exist there – they break down. In stellar interiors there can be only transformation of complex systems, called the nuclei of atoms, once considered indivisible. At temperatures of millions of degrees there is a breakdown not only of atoms but also of their nuclei and a shuffling of decay products from which the nuclei of new elements are formed. Such transitions are called nuclear reactions.

The nuclear theory leads to the conclusion that the energy source of the main-sequence stars, including the Sun, is the continuous formation of helium atoms from hydrogen atoms.

It is well known that the mass of a helium atom is approximately four times that of a hydrogen atom.


Hydrogen envelope (expansion of several million kilometres)
15000 km
5000 km
Helium-rich core
Hydrogen burning shell
Carbon core
Hellium burning shell

However, it takes more than just four hydrogen atoms to have a helium atom: a whole series of remarkable transformations short of magic are necessary. Such transformations result in a release of energy, and hence in a change in mass. So the helium atom receives somewhat less than the mass of the four hydrogen atoms.

Before taking a closer look at the way in which these transformations occur and the helium factory operates in the interiors of stars, it is first necessary mentally to delve into the tiny atomic nuclei.

Nuclei and Nuclear Reactions

As early as 1869 the Russian chemist Mendeleyev arranged all the then known chemical elements in a table that is now known universally as the Periodic table. The elements are arranged in this table in the order of the weights of their atoms and the places of the elements in the table are related to their chemical properties. All the later discovered chemical elements fell into the cells of the table that had been left empty by Mendeleyev and have now reached a total of 107. The sequence number of an element in this table, beginning with 1 for hydrogen and ending with 107, is called the atomic number and is designated Z. Physicists later discovered that there are atoms having a somewhat different atomic weight but completely identical chemical properties and called them *isotopes*. Many chemical elements represent a mixture of such isotopes, the abundances of such isotopes in their natural mixture being almost always constant. The atomic weight of each isotope, denoted A, in comparison with the weight of an atom of hydrogen is almost precisely expressed by a whole number. Different abundances of isotopes are the reason why the periodic table contains atomic weights not expressed by whole numbers. It was discovered, in particular, that there is an isotope of hydrogen with atomic weight 2 (heavy hydrogen or *deuterium*) and an isotope of helium with atomic weight 3, whereas earlier it had been assumed that the atomic weights of these elements were close to 1 and 4. The abundances of "heavy hydrogen" and "light helium" are negligible and therefore the weight of "ordinary atoms" (i.e. the mean atomic weight) of these elements is very close to the weight of their principal isotopes, 1 and 4 respectively. This all changed the former idea that a difference in atomic weight is the principal reason for a difference in the chemical properties of the elements.

The accompanying table gives the atomic weights A and atomic numbers Z of isotopes of the lightest chemical elements and the percentage of each isotope as known on Earth. For example, we see that beryllium has no isotopes but that oxygen has three.

The idea that atoms were small indivisible spheres had to be replaced by a more complex concept. A normal atom with atomic number Z consists of a tiny nucleus (with a diameter of the order of 10^{-13} centimetres), surrounded by a group of Z electrons. An electron carries the minimum possible charge of negative electricity. The charge Z of electrons is equal to the charge of the nucleus, but the latter is positive. If such an atom loses one or two electrons, it acquires a single or double positive electrical charge, becoming a *positive ion*, and if it succeeds in drawing an extra electron into its shell, and thus obtaining a single negative charge, it becomes a *negative ion*. Thus, the neutral hydrogen atom possesses one electron, a helium atom 2 electrons,

The structure of a star during the helium burning stage.

Isotopes of Light Elements

Z	Symbol	Element	A	Percentage
1	H	Hydrogen	1	99.98
			2	0.02
2	He	Helium	3	0.00001
			4	99.99999
3	Li	Lithium	6	7.9
			7	92.1
4	Be	Beryllium	9	100.0
5	B	Boron	10	18.8
			11	81.2
6	C	Carbon	12	98.9
			13	1.1
7	N	Nitrogen	14	99.62
			15	0.38
8	O	Oxygen	16	99.76
			17	0.04
			18	0.20

and so on. The mass of the electrons is 0.000544 of the mass of an atom (for all practical purposes, of the mass of the nucleus) of hydrogen and it would take 1836 electrons on one side of a scales to balance one hydrogen nucleus on the other side. Electron losses exert little influence on the mass of atoms.

The chemist, dealing with hydrogen or nitrogen, denotes them H and N. The physicist distinguishes their isotopes, differing in mass, designating them ^{1}H and ^{2}H, ^{14}N and ^{15}N, placing at the upper left of the chemical symbol the atomic weight A. Having in mind the nuclei of atoms, he places the atomic number to the lower left, e.g. $^{1}_{1}H$, $^{2}_{1}H$, $^{14}_{7}N$, and $^{15}_{7}N$. The electrons are denoted by e^{-}.

1932 started a succession of discoveries of new particles, beginning with the *neutron*, approximately equal in mass to the hydrogen nucleus (called a *proton*, meaning "simplest"), but without an electric charge.

The "simplest", however, is not that simple and even in theory it is impossible to construct from it the nuclei of other atoms. If a nucleus of an atom with mass A consisted merely of A protons, its charge Z would equal A, whereas in actuality A is always greater than Z (except in the case of hydrogen itself). In most cases A is approximately twice as large as Z.

It was shown by the Soviet physicist Ivanenko and the German physicist Heisenberg that the nucleus of an atom consists of Z protons and $A-Z$ neutrons, its charge being Z and its atomic weight $Z+ +(A-Z)=A$.

The negative electron was found to have its counterpart, the *positron*, which had the same mass and size but an equal positive charge. The positron had escaped notice until then because it is short-lived. Under terrestrial conditions in a millionth of a second it collides with an electron and both are transformed into two photons.

In chemical reactions the atoms interact with one another, leaving the nuclei inside unmodified. In nuclear reactions the atomic nuclei are rearranged and form a new atom with completely new chemical properties.

The chemical reactions are expressed by formulas of the type

$$2C + O_2 \rightarrow 2CO,$$

which means that two independent carbon atoms combine with a molecule of oxygen, consisting of two of its atoms, to yield two molecules of carbon monoxide CO.

Similar notation is used to describe nuclear reactions.

The alchemists of old long sought the "philosopher's stone", a method for transforming cheap metals into gold. Their work was in vain and their dreams were not realized, since the atoms in their hands did not wish to be transformed into others. In the 20th century, however, science discovered that with certain atoms this miracle occurs by itself, but gold is not an end product. For example, the atoms of the radioactive elements uranium and thorium undergo a long series of transmutations into other atoms. These unusual transmutations of uranium and thorium atoms are accompanied by the ejection of nuclei of a lighter element – helium – and the appearance of electrons and very short-lived "hard" (with a small wavelength) electromagnetic rays called gamma rays. Eventually, uranium and thorium are transformed into lead.

The nuclei of helium atoms, or α particles (moving at about 20,000 kilometres per second), were the shells scientists were able to use to destroy the nuclei of certain other atoms and discover their nature. Flying towards them at an immense velocity, the α particles smashed these nuclei and formed new chemical elements from the fragments. This required that α particles be available and beamed in the required direction. In 1919 the famous British physicist Ernest Rutherford realized the dream of the alchemists – artificial transformation of the elements.

By causing the collision of a nucleus of $^{4}_{2}He$ with a nuclei of nitrogen $^{14}_{7}N$, he succeeded in transforming them into two other nuclei: hydrogen and oxygen in the form of a rare isotope with atomic weight 17. But as we have seen, oxygen thereby does not cease being oxygen.

This miraculous transmutation can be written using the formula

$$^{14}_{7}N + ^{4}_{2}H \rightarrow ^{17}_{8}O + ^{1}_{1}H.$$

You see that the sums of both the subscripts and superscripts on either side of the equation are equal.

The α particles are supplied by radioactive atoms, but still more "armour-piercing", or rather "nucleus-

piercing", shells are obtained from artificially produced and accelerated protons and nuclei of heavy hydrogen (deuterons).

This is accomplished by the use of powerful accelerators, which accelerate protons, deuterons and electrons to enormous energies, comparable to that of cosmic rays.

Neutrons, which have no charge and are not repulsed by nuclei, penetrate even more easily into the nuclei and have a more destructive effect, as are hard gamma rays.

In recent years a great many artificial nuclear reactions have been produced. Of particular interest are those leading to new radioactive nuclei. The new nuclei are found to be extremely unstable, decaying spontaneously (i. e. not induced externally) and very rapidly, so that they are not encountered in nature. In the long run each element was found to have one or more radioactive isotopes. Most of these "artificial" radioactive nuclei eject only electrons or positrons, rather than emitting α particles.

In the artificial transmutation of the elements associated with the disintegration of nuclei, the destroying particles have a large energy and must move rapidly. They are accelerated to such energies by an intense electric field produced in the laboratory. In nature, however, the energy required for the particles to destroy nuclei comes from high temperatures. As is well known in physics, an increase in gas temperature leads to more vigorous motion of the molecules or atoms making up the gas. Their velocities can be derived from the gas temperature, and thus from the atomic mass we can determine the energy, which goes into the work of destruction of nuclei.

The creators of armour-piercing guns, given the mass of the shell and the thickness of the armour, work out the velocity to be imparted to the shell to break through the armour. Similarly, we can compute the temperature at which the energy of the destroying particles is sufficient for their penetration into atomic nuclei.

For example, two protons in a head-on collision can overcome the mutual repulsion only at a velocity caused by a temperature of 55,000,000 K. Where can such temperatures occur? Not in the laboratory, nor even on the surface of stars. We can only expect to find such temperatures inside the stars, and we arrived at this conclusion long before we began to study nuclear reactions. There, in the mysterious and unseen depths the weight of the above-lying layers of stellar matter creates extraordinary pressure and high gas density. In the hellish conditions of star interiors the particles, travelling at tremendous velocities, collide with one another with the result that the outer parts of the atoms, i. e. their electron shells, are torn off. At these temperatures and pressures the nuclei of all the light atoms should be stripped of electrons, so that in the resultant mass of fragments the number of detached free electrons will be greater than the number of nuclei. Some of them manage to capture an electron as it passes by, but not for long. The next collision returns the atomic nuclei to its solitude. Under terrestrial and laboratory conditions the shells of outer electrons, like a shield, partly protect the nuclei from fatal collisions, but inside the stars the only barrier is the mutual repulsion. The best protected are the nuclei of heavy elements, in which the nuclear charge is large and there is therefore a large force of repulsion.

It is interesting to note the following property of this "crumbled" matter making up the interiors of stars. In the computation of the values characterizing the different physical conditions and events inside the stars an important role is played by the mean atomic weight of the constituent particles. It would appear that it should be strongly dependent on the percentage of different chemical elements because the atomic weight of hydrogen is 1 and of uranium 238. But with the complete ionization of atoms in the interiors of stars each of them breaks down into $Z+1$ particles (1 nucleus and Z electrons). This mixture of atomic fragments will then have the atomic weight $A/(Z+1)$, not A. For example, for pure hydrogen this will be 0.5 and for pure uranium 2.6.

Thus, the paucity of evidence concerning the precise chemical composition of star interiors is of no consequence for the estimates of the mean atomic weight of the particles. There cannot be many heavy atoms in the interior that must be dominated by hydrogen. The data available lead us to assume that at least 50 per cent (by mass) of the Sun consists of hydrogen, and therefore the light-weight hydrogen atoms are the predominant species. The mean

atomic weight inside a star should be close to 1.

To obtain the rate and effect of the nuclear reactions in the gas we must know the atomic structure and the conditions inside the stars. We thus need qualitative data suitable for mathematical calculations. Otherwise we will not be dealing with scientific theory but with assumptions.

To begin with, it is necessary to trace the fate of the particle that penetrates into a nucleus. It appears that sometimes the particle can simply pass through the nucleus. On the other hand, the nucleus can retain the penetrating particle, releasing the energy imparted to it in the form of gamma radiation. Finally, a nucleus that has been penetrated by a particle can break up, as in one of the reactions described above.

Considering the variety of structures of nuclei it is possible to expect a great diversity in the types of collision and their effects; experiments confirm these expectations and show that for each particular type of nuclei certain velocities of collision are more favourable for attainment of the desired result than others. For example, the reactions in which a complex nucleus, forming from two colliding nuclei, breaks into two (by no means equal) parts, are far more probable than reactions in which only an electron or gamma quantum is ejected from the complex nucleus.

The computations of the probability and rate of various nuclear reactions at various temperatures were begun in 1929. We are interested, of course, in those reactions which are accompanied by the release of energy.

In the nuclei of atoms there are surprising transformations which would appear improbable to us if we had not been convinced of them directly. The energy released in these transformations is even difficult to conceive, so immense is it. It has been found that the total energy of any body is related to its mass and this relationship is expressed by the formula

$$E = mc^2,$$

where E is the energy, m is the mass and c is the velocity of light.

In order to describe the magnitude of the energy being released it is sufficient to say that the burning of a tonne of coal in pure oxygen liberates about $5 \cdot 10^{16}$ ergs of energy, whereas the release of all the energy contained in one tonne (it makes no difference whether it is coal, straw or some other substance) would yield 18,000,000,000 times more than this. If we released all the energy associated with a piece of coal the size of a pea, it would supply a huge ocean liner with all the energy needed to circumnavigate the globe. But for the time being we cannot set free all the energy associated with any mass. In particular, the loss of mass associated with the transformation of matter into radiation in radioactive decay is fractions of a per cent.

In radioactive decay and in nuclear reactions energy is released due to a decrease in the mass of the elementary particles participating in the reactions. The end result of such events, occurring with myriads of atoms, yields an appreciable result. For example, we can reliably measure the heat released in the laboratory by radioactive matter in the volume of a thimble. Finally, the solar heat has been sensed by man since the first days of his existence. Solar heat has made possible the origin and evolution of life on Earth and we can say that in a sense we owe our existence to it.

In order that future technology should be able to use the prodigious reserves of energy stored in atoms so that we could replace many tonnes of coal needed for fueling an ocean liner by a test tube containing some substance whose energy would be used to power the ship we will have to travel a long road. Not all the secrets of atomic nuclei have yet been unravelled, a study of celestial bodies infinitely distant from us and apparently of no use to us will be of help. Unfortunately, many people do not suspect that astronomy, regarded by them as a diversion far removed from everyday life, is a highly necessary link in the development of both spiritual and material culture.

Nuclear reactions are the principal source of energy in the stars and the Sun. Decreasing the mass of particles participating in these reactions also decreases the energy associated with them and the liberated energy is radiated into space. Consequently, the solar mass decreases by radiation.

Knowing the intensity of solar radiation, we conclude that every second its mass decreases by 4 million tonnes carried off by radiation.

An extraordinary figure, but it is impossible to

discover such a decrease directly because it is negligibly small in comparison with the mass of the entire Sun. During the time that the Earth's crust has existed (3 milliard years) the Sun has lost only 1/7500 of its mass. If someone could have determined the solar mass with modern accuracy at the birth of our planet, even by comparing it with present-day determinations we still could not determine its decrease. The accuracy of determination of the mass of celestial bodies does not attain 0.01 per cent, i.e. the value that expresses the relative decrease of solar mass during this extraordinarily long period of time.

In nuclear reactions in laboratories the release of energy is accompanied by a decrease in mass not in terms of tonnes, but only fractions of milligrammes, though this constitutes an appreciable fraction of the mass of the atoms themselves. Striving to measure precisely the mass of the atom, physicists have also been able to determine its change as a result of nuclear reactions.

This sort of change shows up in that the masses of atomic nuclei of different chemical elements are not related to one another precisely as whole numbers.

If the atomic weight of oxygen ^{16}O is assumed to be 16, as this is always done, it is found, for example, that the atomic weight of hydrogen is not 1, but 1.00812, the atomic weight of helium ^{4}He is not 4, but 4.00390, and deuterium is not 2, but 2.01470. Let us take a closer look at the latter case. Two hydrogen nuclei, or protons, $^{1}_{1}H$ with a mass of $1.00812 - 0.00054 = 1.00758$ combine to form a pair consisting of a deuteron $^{2}_{1}H$ and a positron. (In each proton there is one positive charge and upon combining they have a double charge; however the deuteron has only one positive charge, like the proton, and the carrier of the excess positive charge – the positron – is released.) This nuclear reaction can be expressed by

$$^{1}_{1}H + {}^{1}_{1}H \rightarrow {}^{2}_{1}H + e^{+}.$$

The positron released in the reaction, whose short life has already been noted, rapidly merges with some free electron. On fusing they both disappear, being transformed into two gamma quanta.

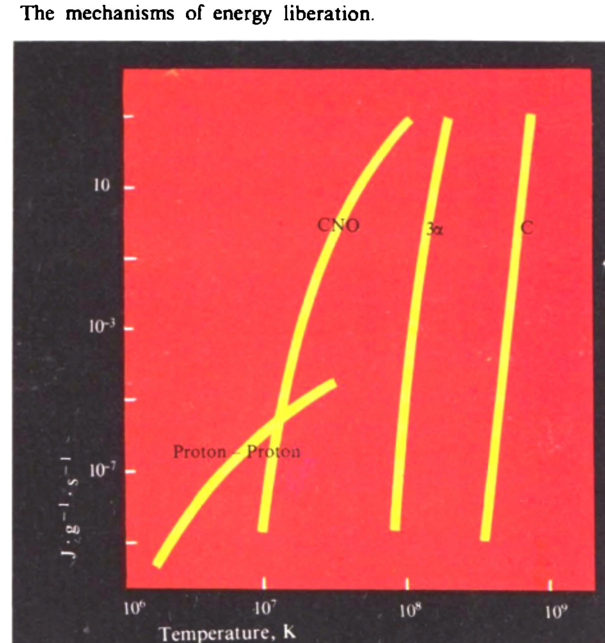
Thus, the joining of two protons and a previously "alien" negative electron leads to the generation of a deuteron and gamma quanta.

The mass of the formed deuteron, together with a positron, is 2.01470, by 0.00046 less than the mass of two protons. Gamma quanta developed from that part constituting 0.00046/2.015 or 0.00022 (almost 0.02 per cent) of the initial mass of the two protons. When a gramme of protons transforms into deuterons, $6.8 \cdot 10^{17}$ ergs are released. This is approximately a thousand times less than the energy in one gramme of protons, but still five times greater than the energy released by one gramme of solar matter during its entire life, i.e. in several milliard years. We see that only nuclear reactions (really existing and observed in laboratory) can play the role of a machine pumping energy to the solar surface.

"Feed Cycle" of Stars

As early as the 1930s astrophysicists were sure that among the nuclear reactions of the light elements the only one capable of maintaining the radiation of the main-sequence stars for a sufficiently long time and with sufficient energy is the formation of helium from hydrogen. Other reactions either continue for

The mechanisms of energy liberation.



too short a time (to be sure, on a cosmic scale), or yield too little energy.

However, the direct fusion of four hydrogen nuclei into a helium nucleus was impossible: the reaction of transformation of hydrogen into helium inside the stars should occur by roundabout paths.

The first path consists in successively combining first two hydrogen atoms and then attaching to them a third, and so on.

The second path involves the transformation of hydrogen into helium with the "assistance" of atoms of nitrogen and especially carbon.

Although it would seem that the first path is simpler, for a rather long time it did not enjoy "due respect" and astrophysicists assumed that the principal reaction providing energy to stars is the second, the "carbon cycle".

Formation of a helium nucleus requires four protons, which of themselves had no desire to merge into an α particle without the assistance of carbon.

In the chain of these reactions carbon plays the role of the necessary helper as a sort of organizer. In chemical reactions there also are helpers of this kind called catalysts.

Thus, in the formation of helium energy is released. In actuality, the chain of transformations was accompanied by the ejection of three gamma quanta and two positrons, also transforming into gamma radiation. The balance is: $10^{-5} \times (4 \times 1.00758 - 4.00390) = 0.02642 \cdot 10^{-5}$ amu.

The energy associated with this mass is set free within a star and makes its way slowly to the surface and then is radiated into space. The helium factory operates in stars continuously until the supplies of raw material (i.e. hydrogen) are exhausted. What happens thereafter will be discussed below.

Carbon, as a catalyst, will last indefinitely.

At 20,000,000 K the effect of the reactions of the carbon cycle is proportional to the 17th power of the temperature. At some distance from the centre of the star, where the temperature is only 10 per cent lower, the energy output decreases by a factor of 5, and where the temperature is lower by a factor of 1.5, the decrease is by a factor of 800. Therefore, no helium is formed from hydrogen at a rather small distance from the central, hottest region. The remaining hydrogen is transformed into helium after the mixing of gases carries it into the area of the "factory", towards the centre of the star.

In the 1950s it was found that at 20,000,000 K, and especially at lower temperatures, the proton-proton reaction is still more effective. This reaction also leads to a loss of hydrogen and the formation of helium. In all probalility it involves the following chain of transformations.

Two protons, upon colliding, emit a positron and light quantum, producing a heavy isotope of hydrogen with atomic weight 2. The latter, after merging with another proton, is transformed into an atom of the light isotope of helium with atomic weight 3, in the process emitting an excess of mass in the form of radiation. If there is a sufficient accumulation of such atoms of light helium, their nuclei on collision form a normal atom of helium with atomic weight 4 and two protons and a quantum of energy into the bargain. Thus, in this process three protons were lost and two appeared - one proton was lost but, on the other hand, energy was radiated three times.

It is apparent that the Sun and the cooler stars of the main sequence of the spectrum-luminosity diagram acquire energy from this source.

When all the hydrogen is transformed into helium the star can still exist due to the transformation of helium into heavier elements. For example, the following processes are possible:

$$^{4}_{2}\mathrm{He} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{8}_{4}\mathrm{Be} + \text{radiation},$$

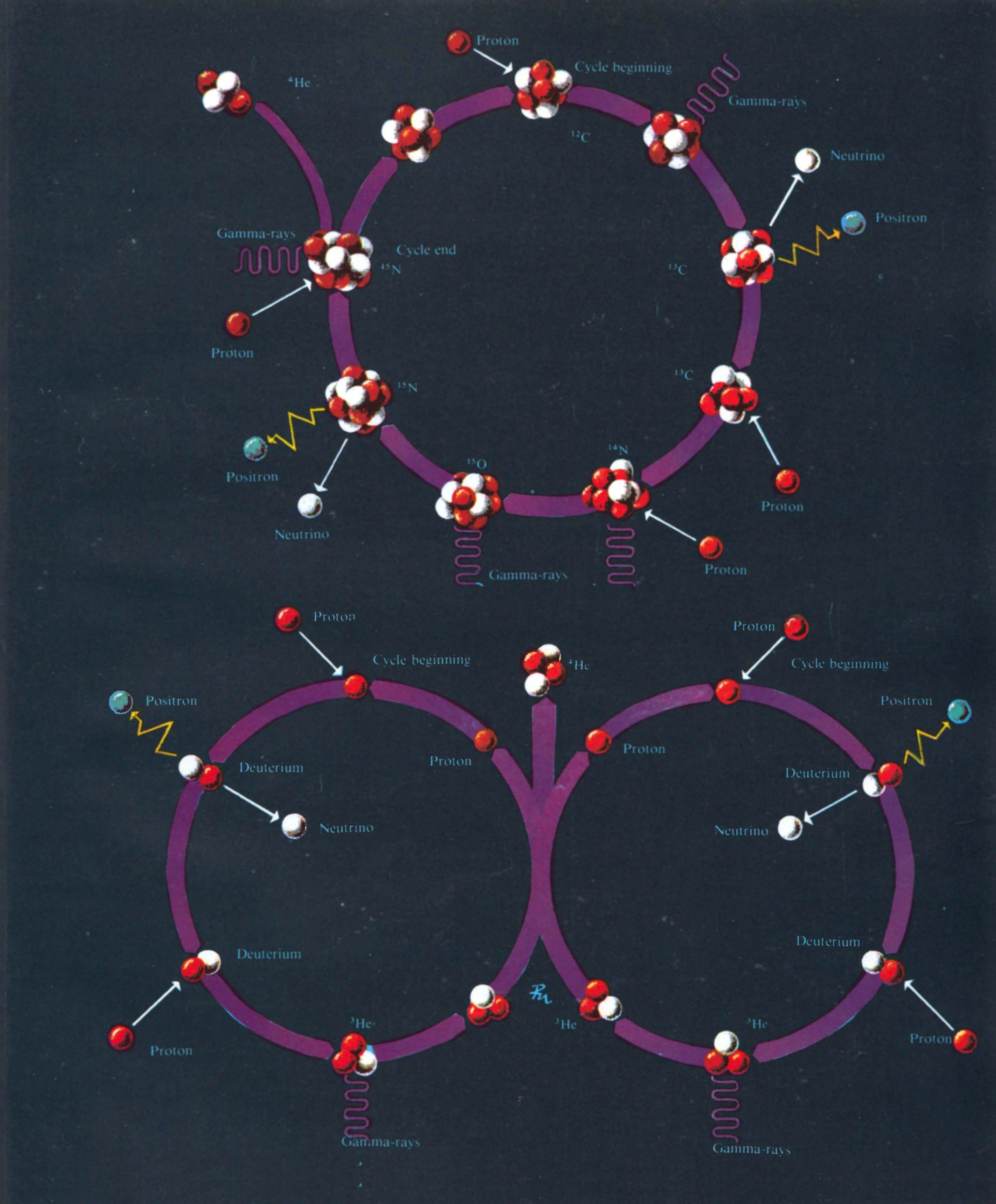
$$^{4}_{2}\mathrm{He} + {}^{8}_{4}\mathrm{Be} \rightarrow {}^{12}_{6}\mathrm{C} + \text{radiation}.$$

One particle of helium gives here 8 times less energy than the same particle in the carbon cycle just described.

Physicists have recently found that in certain stars the physical conditions make possible the development of still heavier elements, such as iron, and they can calculate the proportion of emerging elements in accordance with the abundance of the elements which we find in nature.

In giant stars the mean yield of energy per unit of mass is far greater than in the Sun. However, no generally accepted point of view on the energy sources of red giants has yet appeared. Their energy sources and structure are still unclear to us, but

The carbon-nitrogen cycle and proton-proton reaction.


Proton
Cycle beginning
^{4}He
Gamma-rays
^{12}C
Neutrino
Positron
Gamma-rays
Cycle end
^{15}N
^{13}C
Proton
^{15}N
^{13}C
Positron
^{14}N
^{15}O
Proton
Neutrino
Gamma-rays
Proton
Proton
Proton
Cycle beginning
^{4}He
Cycle beginning
Positron
Positron
Deuterium
Proton
Proton
Deuterium
Neutrino
Neutrino
Deuterium
Deuterium
^{3}He
^{3}He
^{3}He
Proton
Proton
Gamma-rays
Gamma-rays

apparently will be known soon. According to computations made by Sobolev, the red giants can have the same structure and the same energy sources as the hot giants. But they are surrounded by extensive rarefied and cool atmospheres which give them the appearance of "cool giants".

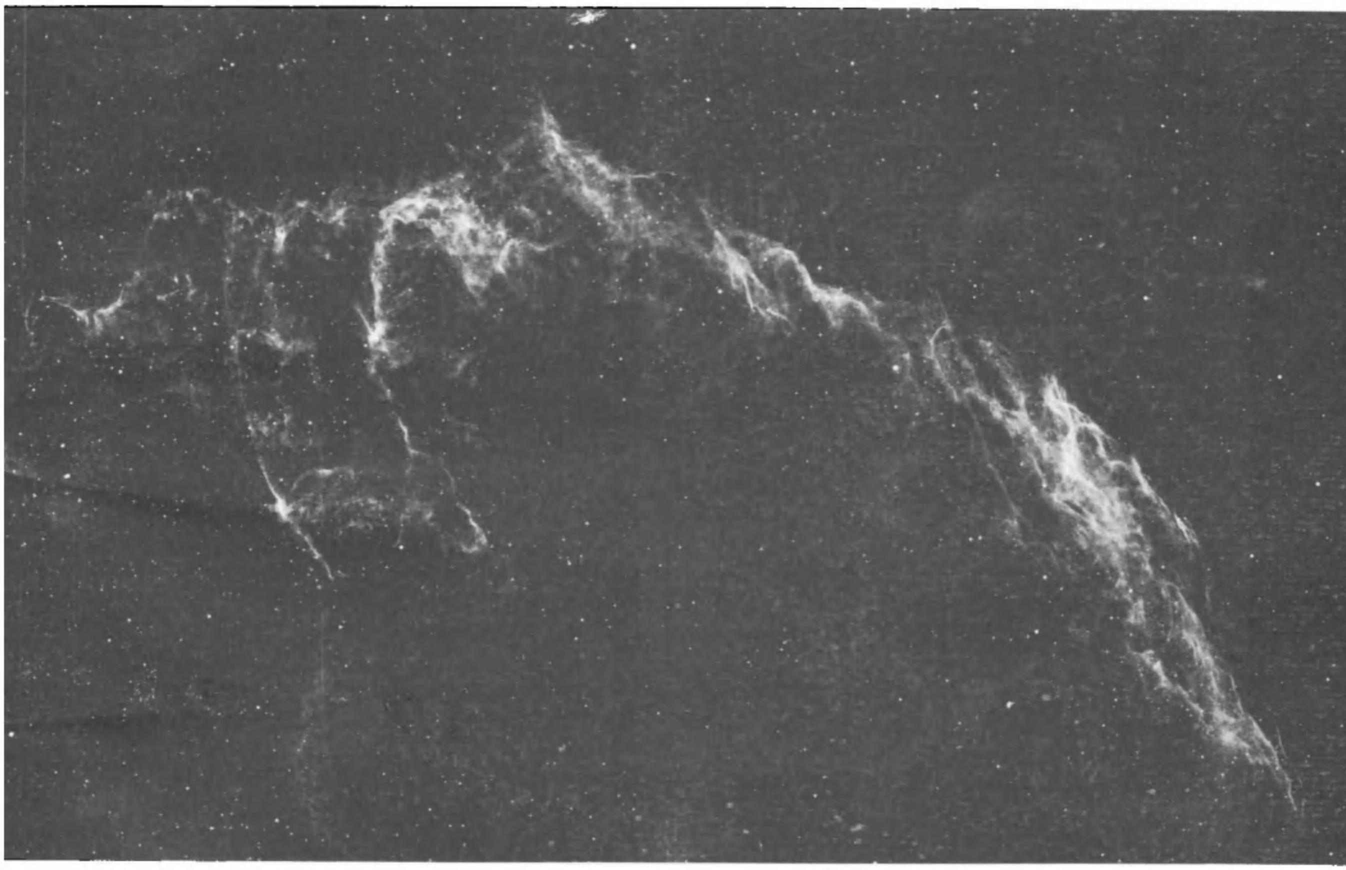
The nuclei of certain heavy atoms can form in stars as a result of combining lighter atoms, and under certain conditions this can occur even in their atmospheres.

Internal Structure of Stars

The origins of the energy sources of stars cannot be explained without a knowledge of the physical structure of star interiors, i. e. without a theory of the internal structure of stars. It is very difficult to check this because the interiors of stars cannot be observed. All that can be done is to construct various theoretical models of stars that are in agreement with the laws of physics and mechanics we know. A comparison of these models with reality can only be done using the relationships between mass, luminosity, energy yield and the properties of the surface layers of stars. But here in turn it is necessary to know the sources of stellar energy. Anyway, Eddington demonstrated in 1916 that without knowing this it would still be possible to form an opinion concerning certain properties of stellar interiors. By 1926 he had developed a theory describing the probable structure of stars, agreeing rather well with observations.

We will view a star as a body subject to various forces. Gravitational pull tends to draw the matter of a star towards its centre, but gas and light pressure from within tend to repulse it from the centre. Since, as we have seen, a star exists as a stable body, without appreciable compression or expansion, there is therefore some equilibrium between the conflicting forces. For this to take place the temperature of the different layers in the star was computed to be such that in each layer the outward energy flux carries to

Parts of a filamentary nebula (NGC 6960/6992).



Parts of an arc-shaped filamentary nebula (NGC 6960/6992) – evidence of a catastrophe of about 160,000 years ago. Today this formation is over 60 light-years across.

the surface all the energy generated underneath it. Two cases can be considered: (1) all the energy is produced at the point centre of the star and (2) the energy is produced by the entire mass of the star but varies in a particular way from layer to layer, depending on its density and temperature. It has been found thus that at the centre of the Sun the temperature is about 13,000000 K and the density is 76 times higher than the density of water. Computations have also revealed that these values change with approach to the surface; other scientists have checked and confirmed these conclusions.

In 1926, when there still was no experimental evidence, it was assumed that there were two energy sources – the release of the energy associated with mass in the combination of a proton with an electron and the formation of helium from hydrogen. The latter was assumed to be more probable.

Twenty years passed and we learned that the truth lies midway between the two above-mentioned stellar models and that they contained far more than a grain of truth.

Helium is in fact formed from hydrogen atoms and their merging is not simple. In this process the mass of the matter being formed decreases because a part of it, together with its associated energy, passes into the mass and energy of gamma quanta. This is the energy source for stars. Finally, energy is actually formed at the centre of a star in a small central core, rather than at a point.

A revision of the theory of the internal structure of stars on the basis of the established source of their

energy changed the above conclusions very little. The history of this problem illustrates once again how the development of science refines the picture of the Universe and how from "obsolete" theories we retain those grains of truth without which ascent to the next step of knowledge would be impossible. In this sense Helmholtz was right when in the last century he postulated that a star derives its energy from compression. We now agree with him, but only for the initial period in the lifetime of a star, when as a result of compression the temperature of its interior increases so greatly that nuclear reactions become possible. Once they have begun they supply it with energy for a very long time (as long as the hydrogen supply in the star suffices). The Sun now contains about 50 per cent hydrogen by mass. This supply will suffice the Sun not only for our generation, but also for many billions of years to come.

One of the small books dealing with the use of solar energy has been somewhat inaptly called "The Sun – an Engine, the Sun – Heat". If we accept this title, it must be remembered that if the Sun is an engine, like any star, it is an "internal combustion engine".

The Origin of Diffuse Matter

Even the ancient Greeks pictured the world as having originated from utter chaos. These ideas of the origin of compact world bodies from tenuous and chaotic matter, usually conceived of as a gas, are unconsciously reflected in the ideas of Herschel about the condensation of nebulae into stars and in the hypotheses of Kant, Laplace, and others concerning the birth of the solar system from nebulae and in Jeans' theory of the formation of spiral stellar systems.

It is difficult to dismiss such concepts because now we can hardly imagine any other process for the formation of stars besides the condensation of rarefied matter into dense bodies. Of the various forms of matter in the Universe at the present time we know, besides the large bodies (planets and stars), only diffuse gases and meteoric dust.

As a corollary to the concept of the condensation of gas into stars, since the times of Herschel diffuse nebulae such as the Orion Nebula were considered remnants of the primary nebula, as a sort of trimmings of matter from which the stars were composed. This concept was not questioned for more than two centuries but a considerable body of evidence has accumulated which permitted me to come out in 1931 with a hypothesis of a completely different nature. It seems to proceed by itself from the sum total of observations.

In short, the presently observed diffuse nebulae and interstellar gas and, perhaps, also the interstellar dust should, for the most part, be considered, a product of the activity of stars. Masses of diffuse gas are forming at the present time, as it were, before our very eyes.

We will first of all point out that a radial expansion with a speed of tens of kilometres per second has been detected at many planetary nebulae. From the cosmic point of view, slowly but steadily planetary nebulae, those gaseous shells surrounding their star nuclei, expand like soap bubbles. A gaseous envelope, which is moving at a speed of tens of kilometres per second and lagging behind its own star by hundreds of astronomic units, cannot be held up by it. The expansion of the nebula, its rarefaction, and dissolving in space are inevitable. Sooner or later small planetary nebula of clear outline deteriorates, turns into interstellar gas, and loses ties with its star. If the mass of a planetary nebula is sufficiently large, after having expanded, still being sufficiently dense, it will occupy a space, such that it will change into a diffuse nebula. A diffuse nebula differs from a platenary one only by its larger dimensions and irregular form. But as it is highly unlikely that all parts having a different density will expand absolutely symmetrically, the planetary nebula should become more and more irregular with the passage of time.

In the sky you can see examples of nebulae of the type in transition from planetary to diffuse nebulae, and in general the greater the dimensions of a planetary nebula, the more it approximates the diffuse nebula type.

There are absolutely no doubts that planetary nebulae are formed from gases released at some time by star itself. And these gases supply the bulk of thin interstellar gases, and even diffuse nebulae.

Already in a hundred million years gases of an expanding planetary nebula completly lose their link

to the parent star and come within the range of action of other stars. I estimated the mass of the envelope of a planetary nebula to range from 0.1 to 0.01 of the solar mass, and the Galaxy at present contains many thousands of such nebulae. Assuming that in the Galaxy there always existed only ten thousand planetary nebulae simultaneously and that the Galaxy is just as old as the Earth's crust (and this is the least possible age of the Galaxy), we come to the following conclusion.

During the lifetime of the Galaxy, stars that form planetary nebulae around themselves have been supplying a mass of gas into outer space equal at least to the mass of ten million suns – an impressive mass, which probably is many times greater.

Besides planetary nebulae, novae and supernovae throw out gases directly before our eyes into outer space. Even if we leave aside the supernovae, which are as yet poorly studied and which throw off large masses of gas, even then the mass given off by normal novae is sifficiently impressive.

A supernova explosion ejects from 10^{-4} to 10^{-5} solar masses, and dozens explosions take place in our Galaxy every year. If during the lifetime of the Earth's crust the novae in the Galaxy erupted as frequently as now, then during this time they emitted into interstellar space as much gas as was supplied by the planetary nebulae. Apparently, the supernova explosions give just as much.

In a year, a Wolf-Rayet star loses about 10^{-5} solar masses by continuously throwing off atoms from its surface. Apparently, this process continues for about ten thousand years. Two different methods for estimating the number of Wolf-Rayet stars in the Galaxy give the same number of four hundred thousand. If such a proportion existed in the Galaxy since the Earth's origin, then during this time interstellar space acquired a mass of gas from which it would be possible to make 3,000,000,000 suns.

If the Wolf-Rayet stars are capable of throwing off atoms so energetically for only ten, not ten thousand years, then as suppliers of gas to space they would not be inferior to the planetary nebulae and novae. Even our Sun and all the stars lose some matter from their surface, filling the surrounding space with it.

If we take into account the fact that all the stars supply their gas to interstellar space (by prominences or otherwise), then we will find that the amount of gas produced by the stars during the life of the Galaxy may even be larger than the amount of diffuse matter observed in it now. It follows then that the diffuse matter has a sink. Where can it go? It seems to recondense into denser bodies such as stars.

Thus, all the above-mentioned stars, which are sources of gas, during a time undoubtedly less than the age of the Galaxy, have scattered in space a mass of gas from which it would be possible to make billions of suns.

Various ways of estimating the amount of diffuse matter in the Galaxy (both in the form of gas and in the form of dust) lead to values between 10^8 and 10^{10} solar masses. The amount of gas thrown off by the stars is thus sufficient to form all the existing gas nebulae and the interstellar gas atmosphere and even the dust nebulae (light and dark).

It is generally accepted now that stars eject huge amounts of gas, which again goes to form stars. It is, of course, necessary that the first generation of stars in each galaxy be formed from the "primary" gas, which comes from some other source, rather than from stars.

It should not be thought that the rotation of gas and stars noted by us causes an endless repetition of what has passed. As with the rotation of life on Earth, rotation in the Universe leads to changes. Plants and animals now on Earth are not the same as those which existed on Earth millions of years ago. Stellar emission is accompanied by the conversion of hydrogen into helium, and in the stars heavy chemical elements seem to form. Therefore, the chemical composition of stars changes and, for this reason, the composition of the gases, which they throw off and from which other stars then emerge, changes. The composition of stars being born is now different from that of stars born earlier, and not all stellar matter is dispersed in space. The amount of gas in the stellar systems dwindles. Observations show that there are stars rich in hydrogen or metals, or else poor in them. Apparently, these are stars of different periods of formation, condensed from gas of different chemical composition, and not always or differently "cooked" inside the stars.

By contrast, Ambartsumyan believes that nebulae and hot stars originate simultaneously from some

superdense form of matter, and that for the most part nebulae form not by gas ejections from stars, although he does not deny the possibility of such ejections of substantial masses.

The question arises–could the gases have formed interstellar cosmic dust? Something has been clarified on this subject only quite recently. Using conclusions concerning the process of condensation of metal vapours on solid bodies, some scientists consider that the condensation of gas molecules into particles of cosmic dust is possible due to the difference in temperatures of these particles and interstellar gas. The energy of the collision of an atom with a dust particle is quickly radiated into space or partially changes into internal energy of the atom so that the dust particle remains cold. For a milliard years the mass of particles, which initially, perhaps, is extremely small, reaches 10^{-15} grammes, as calculations show, and this is just the mass of particles of dark nebulae.

Tiny condensation centres may emerge during the breakdown of bodies such as a cometary nucleus. It is sufficient for a negligible part of all cosmic dust to have such an origin, and this will assure the further growth of each dust particle.

A certain percentage of slow atoms may also enter directly into molecular combinations and then into groups of molecules, i.e. centres of dust particles.

Where gases ejected by the stars gather into clouds and thus form diffuse nebulae, they may shine if there is rather a hot star to stimulate their glowing. When the nebula has accumulated enough particles, a dark dust nebula is formed that shines when nearby there is a bright, if not hot, star. Several scientists developed the theory that meteorite dust particles grow by sticking together.

The Origins of Stars

Mounting evidence is available that all the stars, even in the Galaxy, cannot have originated at the same time. Undoubtedly, among the stars there are older and younger ones.

It may well be that somewhere in the Galaxy stars are being born at present. The naked-eye stars are for the most part not kids or elders. Our stellar system as a whole should be viewed as being in its prime, although it, of course, includes both elders and new-borns, and we only now begin to learn to distinguish them. The consistent suggestion that stars are born in the Galaxy even at the present time was made in 1947 by Ambartsumyan. The hypothesis stimulated a wide spectrum of studies to confirm that age differs from star to star and from stellar system to stellar system. Another of his pioneering ideas was that groups of stars can be born from some superdense bodies. Since then superdense bodies, neutron stars, have been found in nature.

Convincing evidence of the young age of stars is the cases where they are observed arranged in a line, chain or group–association–but not so close that it could be held together by mutual pull. Of especial interest is the discovery made by Herbig of the Lick Observatory. He found in Orion some very faint stars, each of which is surrounded by a tiny nebular envelope with unusual glowing. These stars are fairly cool and resemble T Tauri irregular variables showing low luminosity. In 1954 Herbig found in that close group of nebular stars two new stars of the same type, which were not there seven years before. Since then these small stars do not change their brightness. This perhaps was the birth of stars; they with time will become T Tauri variables, which are often associated with unusual faint nebulae.

Can it be that stars are condensed from interstellar dust?

Normal dust nebulae show no clear signs of their matter turning into stars. Recently very small dust nebulae have been found, which, however, absorb light appreciably. They were called *globules*, and a relatively small number of them is only seen against the background of several light and large nebulae. Some people believe that further contraction of globules, which, of course, contain gas as well, heats them and makes them glow and turn into gaseous stars. But it is unclear where the great amount of hydrogen in stars that originated in this way from interstellar dust comes from. It is also sometimes assumed that among interstellar dust there should be much frozen gases containing hydrogen, which is so abundant in the Universe.

Most astronomers believe that the largest and hottest stars belong to the youngest stars. They waste their energy so lavishly and generously that they would have long ago gone bankrupt, if they had emerged long ago. If we still see them now as they

waste their energy so, this suggests that they became wasteful only recently, no more than 10 million years ago. And we do not know other stars with the same, or larger, mass, but more thrifty, that could be treated as being in the stage just before hot giants. Some of variables are young too.

Ambartsumyan supposes that the apparent clustering of giant hot stars exactly reflects their clustering in space. These groups of hot stars, which he called O-associations, are thought to be systems intermediate in terms of size between stellar clusters and stellar clouds, unstable systems at that. He also supposes that the hot gases of which they are formed have emerged here fairly recently from some superdense prestellar matter. After their group ejection they fly apart and, spreading in space, already in a million years turn into cooler stars.

Ambartsumyan thinks that stars of fainter luminosities emerge as T Tauri irregular variables, and he called their clusterings T-associations, assuming that these groups likewise expand, having formed from some superdense prestellar bodies.

Based on our studies, we question the existence of radially spreading groups of hot giants of the above size. Admitting that these stars are young, we believe that they only enter normal stellar clusters and huge stellar clouds, forming relatively small groups like fluctuations. But the apparent clustering of such stars we explain by the fact that in the neighbourhood, within the Milky Way, such stars, just like other distant stars, are blotted out by cosmic dust clouds.

In the gaps between the clouds so-called corridors of visibility emerge. We see in them hot stars and other objects lying at various distances and projecting one upon another, hence the picture of their apparent clustering. To be sure, the hot stars, just like all the other ones, are distributed over clouds nonuniformly, but this is another story.

My studies have shown that narrow and long filaments, often arranged as rectangular links on the branches of spiral arms in spiral stellar systems, are the most likely clouds where hot giants and other stars in spiral galaxies are born. Here tens and hundreds of giants enveloped by the gaseous nebulae produced by the giants, are arranged as bunches of grapes. Most of open clusters must be born here. The new-born giants and other stars tend to spread out, having different velocities at birth. As a result, narrow and bright spiral arms gradually turn into large clouds consisting, in particular, of hot giants, into the clouds whose spiral arrangement becomes less obvious. In the process, giants and other stars continue to be born both in former places and, as an exception, in detached fragments of spiral arms.

The Life and Death of Stars

The current theory of evolution, i.e. life, of stars is based on the theory of their internal structure and sources of stellar energy. Also, it is based on physical theories, such as thermodynamics, hydrodynamics, nuclear physics, radiation transfer theory, etc., and it requires advanced mathematics to arrive at numerical results. Beginning in the 1950s, with the advent of computers, computations became much easier.

Having originated as a thickening in dust-gaseous medium, a star's only energy source is gravitational contraction, until the temperature in the centre has reached a value at which the thermonuclear reaction of fusion of hydrogen into helium begins. It has long been calculated that in massive stars this stage lasts hundreds of thousands of years, and in stars with a mass less than the Sun's this stage lasts hundreds of millions of years.

When the temperature of the central region is already sufficient for the energy produced in nuclear reactions to compensate for the surface cooling of the star, its compression discontinues. This equilibrium of heat inputs and outputs occurs at a temperature that is the higher the larger the mass of the star. The energy yield of nuclear reactions, as said above, is strongly dependent on temperature. This accounts for the observed growth of the star's luminosity with its mass (or rather the main mass of the star).

It is important for what follows to know whether or not the mass continues to be sufficiently constant and whether or not stirring (convection) occurs in it, in which the fuel (hydrogen) keeps coming from the outskirts to the core where it burns down.

In 1942 Chandrasekhar and Shoenberg made an important advance. Having assumed, with good reason, as is also assumed now, that a star's mass is constant and there is not stirring, they concluded that the burning of hydrogen in the centre results in the formation there of an ever growing helium core. The star's luminosity must increase 2.5 times by the epoch

when the mass of the helium core reaches 10 per cent of the total mass. Hydrogen takes long to burn out: in massive stars hundreds of thousands of years, in stars with the mass of the Sun several milliard years. That is why it is in this stage that we observe most stars.

An important and unexpected break-through was the theoretical result obtained by Schwarzschild and Sandage in 1952. They found that the helium core, having lost its energy sources, will contract, while the outer layers will expand. Energy will only come from the thin hydrogen layer around the core. As the temperature reaches a certain level, a reaction will begin in the core in which three helium nuclei turn into a hydrogen nucleus, and this added energy feeds the star, which becomes a red giant (or supergiant). The process occurs the faster the larger the star's mass.

To compare the theoretical prediction with observations we will turn to the Hertzsprung-Russell diagram, or spectrum- (or colour-, or surface temperature-) luminosity diagram. We see that the diagram became much more complex than before. Studies of brightness and colour of stars in various open and globular clusters contributed to the theory of stellar evolution. The figures give the famous, classical summary diagram compiled by Sandage in 1957 from observations of a number of open clusters, whose names are provided.

We see that all the clusters have in the lower part of the diagram the main sequence of stars. But the upper branches of the sequence in each cluster extend differently, turning to the right at different magnitude (different luminosity) and different values of colour index.

It is worth mentioning, for example, that the clusters χ and h in Perseus, which include very bright blue stars (at the top left), also have a branch of red supergiants (at the top right), and that the brightest blue stars of the Pleiades are less bright, and in NGC 752 they are faint and yellow, not blue. The right branch in M 67 markedly differs from the others and more resembles the branch for the globular cluster M 3.

This all is extremely important because the stars of one cluster occupy a small volume and they have emerged from the same gas cloud, and therefore must have the same initial chemical composition. The age of the stars in the cluster must be about the same.

Astronomers should be grateful to nature for the existence of stellar clusters, they are just like laboratories where physicists create necessary conditions and exclude the influence on matter of too many factors at once.

On the other hand, the position of the main sequence for the new-born stars (zero-age line or the line of the initial main sequence) and of the line for 1 to 5 milliard year old stars were theoretically predicted. Most of the nearby stars then appear to be younger than 5 milliard years, since in the Hertzsprung-Russel diagram they lie to the left of the curve for $5 \cdot 10^9$ years.

In the figures, the position of stars in the same cluster must only depend on their masses, and the positions of the curves for different clusters must depend on the age and initial chemical composition.

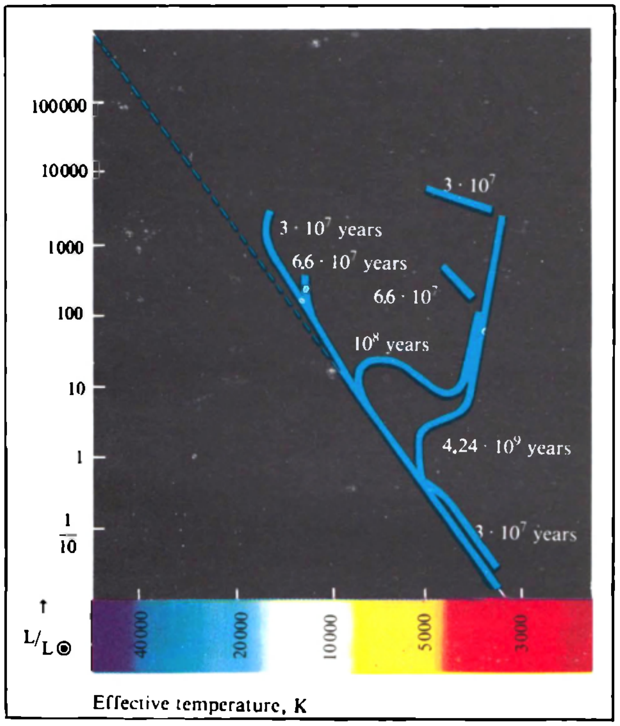
The larger and brighter stars burn down their hydrogen faster, and their life is shorter, whereas sun-like stars remain on the main sequence for about 5 milliard years and the stars 10 times as large for a 1,000 times shorter period. This explains the difference in height of the upper ends of the main sequence for different stellar clusters.

The stars of a very young cluster lie on the main sequence. With time the larger stars first leave the main sequence and shift along the diagram to the right. Accordingly, as a cluster becomes older, the upper level of the main sequence lowers gradually. This age is determined by the time a star of the upper end spends on the main sequence. In other words, it is determined by the position of the point where stars begin to deflect to the right from the main sequence.

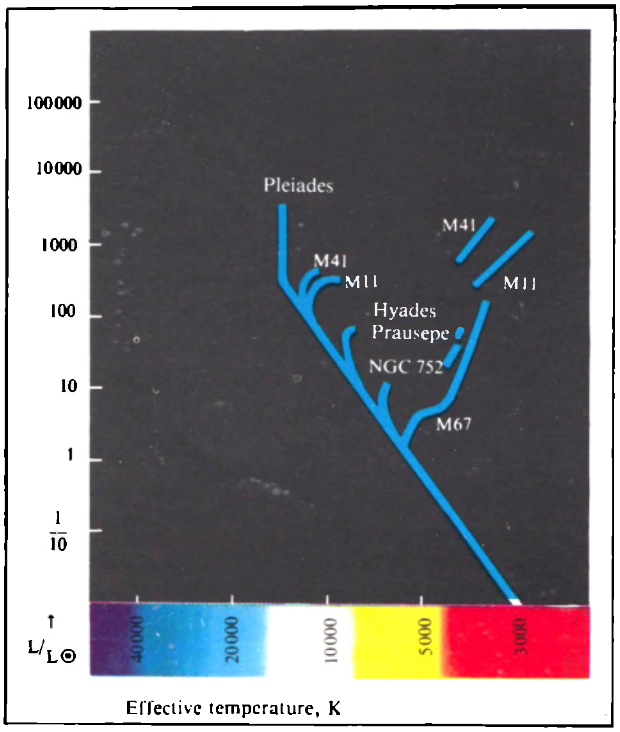
It is found, in such a way, that the cluster NGC 2362 is younger than a million years. The Pleiades are about 20,000,000 years old, and M 67 and M 3 are more than 10,000,000,000 years old.

After the stars have left the main sequence, they, in theory, shift to the right into the region of red giants or supergiants, depending on their mass. In each cluster, red giants or supergiants have the same luminosities as the stars about to leave the main sequence and shift to the right. This corresponds to the development of a helium-rich core in a star instead of a hydrogen core.

But there is a gap between the shifted end of the main sequence and the giants, which is called the Hertzsprung gap after the scientist who was the first to notice it. The gap is large in young open clusters with



The theoretical Hertzsprung-Russell diagram of a hypothetical stellar cluster.



The Hertzsprung-Russell diagram for several open stellar clusters of various ages.

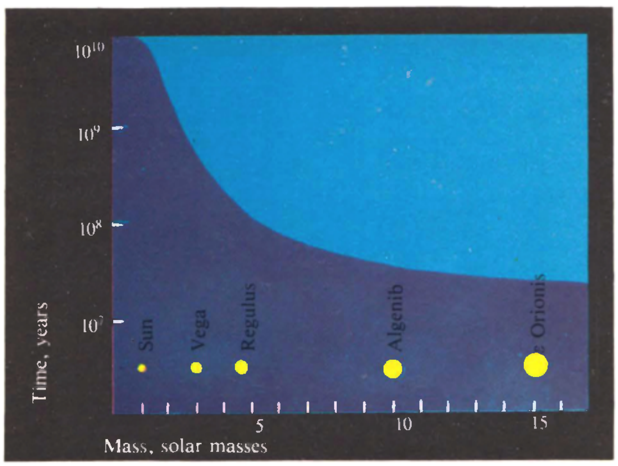
hot stars, and the narrower the older the open cluster and the cooler and fainter its brightest stars. This is because massive bright stars quickly become red giants and therefore it is difficult to find them in an intermediate state. Smaller stars are slower to go over to this state, and so for them the Hertzsprung gap contracts. The globular cluster M 3 has no gap at all, but stars in it begin to deviate from the main sequence at absolute magnitude $+4^m$, i.e. when they are only twice as bright as the Sun.

The difference in the initial chemical composition shows up in the following. The Hertzsprung-Russel diagrams in the globular cluster M 3 and very old open cluster M 67 are very similar and they are about the same age, since their main sequences end at the same point, near $M = +4^m$. But in M 67 the red giants have the brightness 1/10 that of M 3.

Quantitative chemical analysis from spectra indicates that the stars in the halo of the Milky Way and the globular clusters are 100-500 times poorer in metals than the stars from the galactic disk. This makes their atmospheres more transparent. Their radiation comes to us from the deeper and hotter layers and they are whiter and brighter than the stars rich in metals and lying in the same region of the diagram.

It thus turns out that open clusters and stars in them had continuously been formed for 10 milliard years, whereas globular clusters and stars in the galactic halo have all emerged earlier, more than 10 milliard years ago. (Some estimate their age even at 10^{11}-10^{13} years.)

Such a relation between the age and position of the stars in the Galaxy is indicative that when the Galaxy was young, stars were emerging throughout its spherical volume and the gas from which they condensed was distributed in the same way. Later on rotation oblated the galactic gas so that it



The time spent by some stars on the main sequence of the Hertzsprung-Russell diagram.

concentrated near the galactic plane and turned into a disk, where stars continued to form, whereas the galactic halo became depleted.

Before stars, the Galaxy was gaseous and was nearly completely composed of hydrogen. The heavier elements could only emerge in nuclear reactions in stars and they were brought up by convection. The ejection of gases from their surface, especially in catastrophic explosions, enriched the galactic gases by heavy elements. Therefore, the stars that emerged later on in the disk contain more heavy metals.

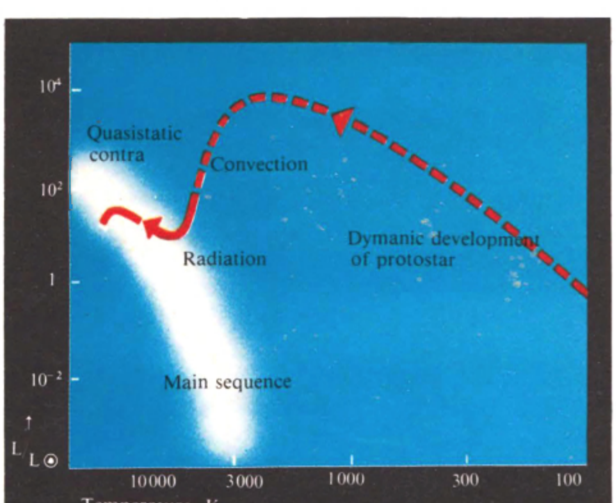
It is now easier to understand the Herzsprung-Russel diagram of the youngest clusters, such as M 16 (or NGC 6611), that are only 200,000 years of age, i.e. younger than the human race. On the main sequence of this cluster lie hot stars of classes from about A0 to O5, and the fainter and cooler ones lie to the right, above the zero main sequence. But these are very young stars, still in the stage of gravitational contraction. According to the Japanese astronomer Hayashi, it is possible to work out the time taken by a star to contract down to a radius and luminosity corresponding to a given sequence. Among such young stars there are many RW Aurigae variables and stars with bright lines in their spectra. These are regarded as suggestive of instability occurring in gravitational contraction.

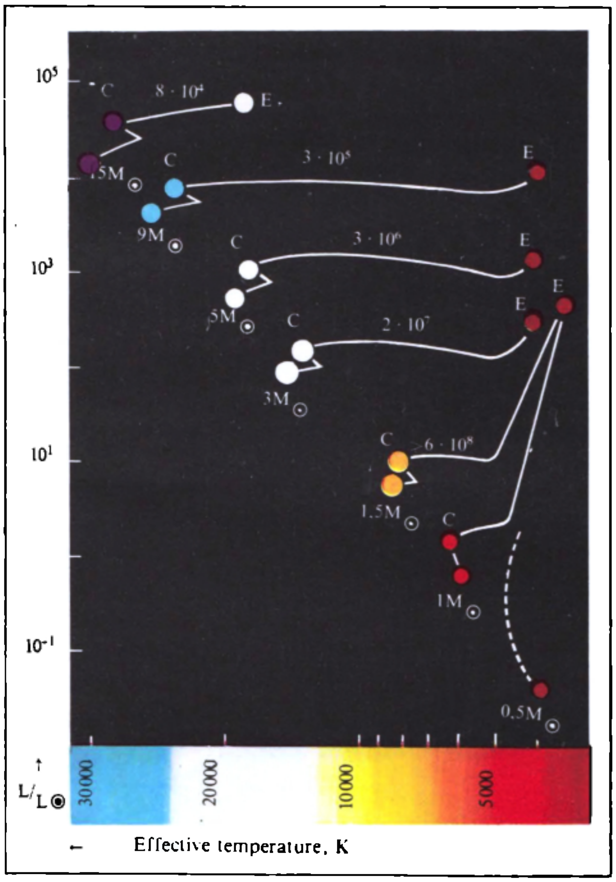
In the stage of red giants, when in a star's core helium is burned out, the star shifts to the left in the diagram, forming a less steep sequence. According to Kipenchan, this motion is complex, with temporal returns, the star evolving with different rate in different sections of the path. In certain regions they for a time even become pulsating stars, Cepheids of different periods. More massive, once bright stars of the spectral class B become long-period variables of large luminosity, and the less massive become short-period variables, especially characteristic of some globular cluster, and with periods less than a day. Their luminosity is lower. A Cepheid is known whose variability nearly stopped. Cepheids fill in the gaps in the diagram, where there are no normal stars. These are the regions of instability and evolution of stars.

Not everything is clear with the Hertzsprung-Russel diagram yet. The latest stages of stellar evolution are obscure, for example. It is, in particular, supposed that having exhausted all the helium, the star contracts quickly – a star in this stage is very difficult to find. It has no more energy sources and turns to an exceedingly dense white dwarf. A white dwarf consumes so little energy that in that stage it can live many billions of years and, as astronomers used to joke, is a "hot corpse". It is unclear whether a star can contract more than a white dwarf. Some believe that it can turn into a neutron star.

But only stars with mass less than 1.4 solar masses may turn into a white dwarf. A massive dwarf is

The path of a protostar in the Hertzsprung-Russell diagram.





The path of stars in the Hertzsprung-Russell diagram after the main sequence stage (the numbers over the curves are the time in years from the end of the hydrogen cycle to the beginning of the helium cycle).

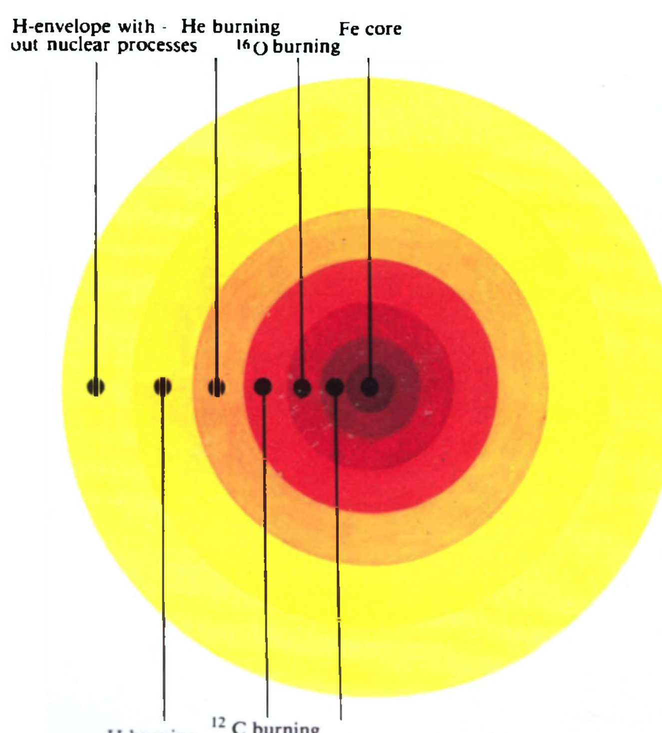
unstable and, maybe, explodes as a supernova, which would be the undoing of massive stars. On the other hand, they may explode repeatedly like novae and, thereby ejecting excessive mass, they also become white dwarfs.

It is to be noted that we as yet do not know any "extinct" stars. The coolest of the known stars, infrared stars, cannot be fading stars. Indications are that they will heat up again.

All the above gives an idea of the current understanding of the emergence of galaxies. Perhaps formerly the Metagalaxy was a huge thickening of hydrogen clouds, which were simultaneously disintegrating into smaller clouds. The clouds flied apart with a velocity that fell off as they receded. Inhomogeneities in clouds produced gravitational condensation of gases into stars within spheres.

As rotation sped up, the gaseous mass enriched by heavy elements coming from old stars flattened a bit. So emerged contracted galaxies with their disks, where younger stars that were richer in metals were born. It is quite probable that the gas in the disk concentrated, with the help of magnetic field, along spiral arms originated from the core, where stars were formed especially intensively and the process carries on now, the more noticeably the more gas is there. Such are spirals of "later types" and irregular galaxies. In the latter, like in the Large Magellanic Cloud, there are young globular clusters, whose Hertzsprung-Russel diagram is rather like that of open clusters. The old open clusters of the Galaxy

The schematic structure of a star after the nuclear energy source in the centre is out.



with diagrams similar to the diagrams of old globular clusters, lie far from the galactic plane. There they were less destructed under the pull of passing stars and, being rich in stars themselves, they were more stable. Their stars, coming apart, join the stellar population of the disk, and the gaseous ejections by the stars supply material for fresh, although fading, stellar formation. But we do not know much about the development of the worlds.

The development and cycles here do not, of course, represent indefinite repetition of the path just covered. But this development occurs in a dialectical way, in the struggle of contradictions, not infrequently by jumps.

During the time negligible in comparison with the cycles of development of celestial bodies human knowledge penetrated the mysteries of their structure and development.

7. The History of the Earth and Planets

A host of various hypotheses are devoted to the origins and history of the planets of the solar system. It is as yet difficult to give preference to any of them. In order that the reader might form some idea of the problem we will recount the hypothesis due to Academician Otto Shmidt, which appears to be one of the most developed theories.

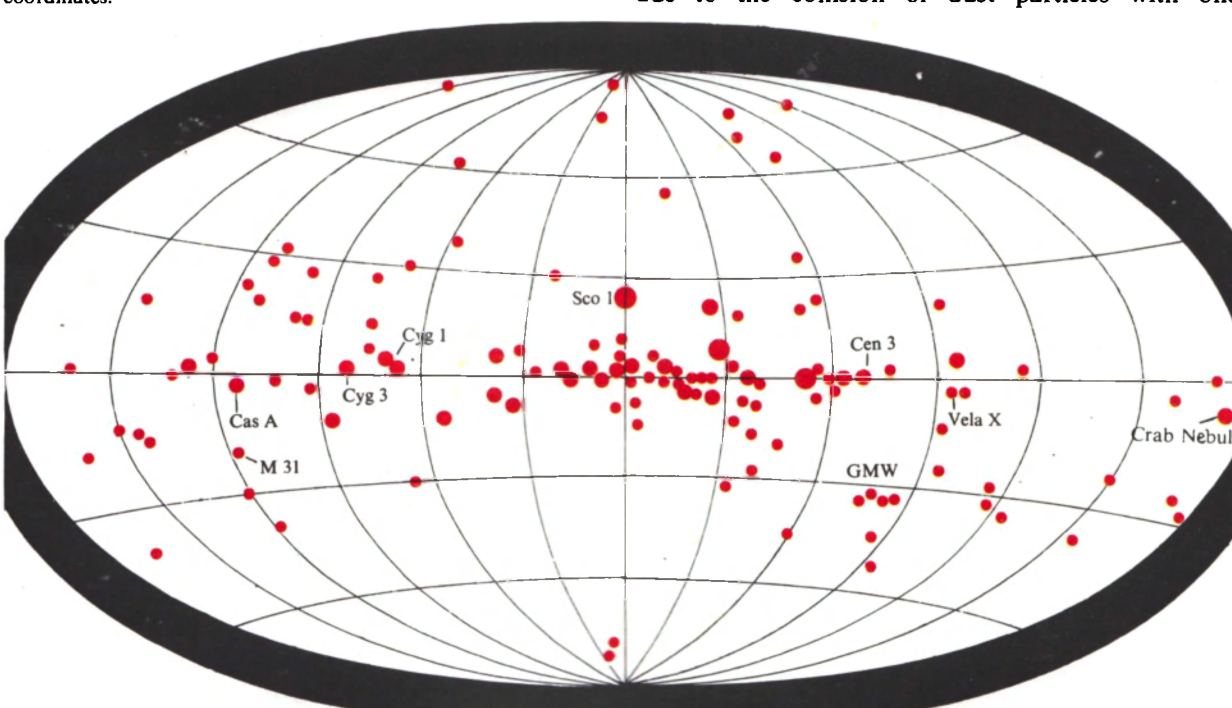
Shmidt first proceeded from the fact that meteoric matter, both in the form of more or less large pieces and in the form of dust, is encountered in the Universe in abundance. Not too long ago, this matter was known to us only within the limits of the solar system, but now we discover it in tremendous quantities in interstellar space. For the most part, meteoric matter is collected in colossal cosmic clouds – in the diffuse light and dark nebulae, which also contain much gas.

Subsequently, various concepts led the Soviet scientists Gurevich and Lebedinsky to the conclusion that pre-planet matter was composed of gas and dust.

The gas-and-dust clouds together with the stars fill the Galaxy, their matter being heavily concentrated towards its plane of symmetry, towards the plane of the Galaxy's equator. Together with the stars, the gas-and-dust clouds participate in the rotation of the Galaxy around its axis. In addition, both the stars and the clouds now approach one another, now recede from one another. Sometimes, a star plunges for a time into the gas-and-dust nebula and blazes a path for itself like a wayfarer in a dense fog. As with the fog for the wayfarer, the gas-and-dust cloud is not an obstacle for the movement of the star; it is not necessary to go astray since its path in the nebula is directed by the same laws of gravity.

Many dust particles fall on a star in the course of its slipping through the nebula, and others, having changed their orbits as a result of the mighty attraction of the star, may be captured by it and made its satellites. Such a capture, however, only occurs under favourable conditions – a reduction in the relative speed of the dust particle due to the attraction of a near-by star or, as Agekyan showed, due to the collision of dust particles with one

The distribution of X-ray sources over the galactic coordinates.



another. In such a "successful" case, the majority of these "acquired" star satellites, according to Shmidt's hypothesis, does not leave the star even after it has left the nebula. The star turns out to be surrounded by a tremendous cloud of gas and dust particles, which describe various orbits around it. Later, Shmidt considered that more probable are clouds from the same diffuse medium from which the Sun itself arose.

The cloud, having been generated around the star, gradually acquired a lens-like form. The rotation of tiny particles in it around the star took place primarily, though not exclusively, in one particular direction (at small angles to one another), and that is why the dust layer pierced by the star cannot be absolutely homogeneous.

In such a star, surrounded by a lens-like gas-and-dust cloud, Shmidt saw our Sun before the formation of the planets.

Of course, not only our Sun could undergo such a meeting with a gas-and-dust nebula. A great number of stars should have experienced such an adventure, and it is in stock for others. That is just as well, because besides our solar system, there should be in the Galaxy an even greater number of planet systems. This inevitable conclusion from a new theory distinguishes it from many other cosmogonic hypotheses, in which the appearance of the solar system is a rare phenomenon.

In the crowd of dust particles rotating around the Sun in intersecting and variously extended orbits, collisions unavoidably took place. And so their movements averaged out and approached circular planes, which lay close to one another. This produced a disk around the Sun from the gas-and-dust cloud, which became thinner and thinner but made up for it by becoming denser. This dense layer of particles, in the parts near the Sun, absorbed its heat. Therefore, farther away from the Sun it was very cold inside the disk, and the gases there froze onto dust particles. This explains why planets farther from the Sun are richer in gas than those close to it. This idea, along with the theory of evolution of a cloud, was developed by Gurevich and Lebedinsky, and Shmidt found that their picture of the cloud evolution was more probable than he showed earlier. The picture of the cloud evolution, although containing a number of additional hypotheses, can be called a theory that lies within the framework of Shmidt's hypotheses. The main Shmidt's hypothesis is the assumption that the planets arose from a cold cloud of particles, and the assumption that the cloud was captured by the Sun when the latter had already been completely formed.

A further picture of the evolution of the gas-and-dust disk is briefly as follows. Dust thickening arose in the condensed cloud and the collisions of the dust particles led to their merging into solid bodies the size of modern asteroids. A great number of them collided and split, but the larger of them, "embryo" planets, survived and "sucked in" surrounding fragments and remnants of the dust, first joining them together by collisions and then, to a greater and greater degree, by attraction. The compact embryo planets were surrounded by clusters of bodies and their fragments, which rotated around them and which gave birth to the satellite of the planets.

From the lens-like form of the nebula surrounding the Sun and from the predomination of parallel movements in one direction, the basic characteristics of the structure of the solar system follow: the rotation of all planets near the Sun in the same direction, small angles between the planes of their orbits, and also the almost circular shape of the orbits.

In one of his first papers, Shmidt computed the rate with which the mass of a planet would grow due to the falling of meteorites on it if the meteorites presently observed in the solar system were remnants of the swarm that earlier surrounded the Sun. It turned out that at first the planet grew swiftly, and then, slower and slower. Roughly speaking, all the "bricks" went into the building of the Earth – bodies of asteroid dimensions and their fragments, which filled the space between the boundaries lying in the centre between the orbits of the Earth and Venus and between the orbits of Mars and Earth closer to the latter.

Of course, it is impossible to determine " when the first stone was laid", but Shmidt's theory permitted computing the time it took the mass of the Earth to double and reach its present value. This time of "semiformation" is close to what may be called the age of the Earth. In any case, this period is a little less than the age of the Earth.

Assuming that *now* more than 1,000 tonnes of meteorite matter falls on the Earth annually, Shmidt worked out the *time of semiformation* of the Earth to be about 7 milliard years. This result is close (on an astronomic scale) to the *age* of the Earth's crust – 3 milliard years – determined by radioactive dating. It is clear that the age of the Earth's crust should be less than the age of the Earth as a whole.

However, since modern meteorites in the solar system are possible fragments of planets located between Mars and Jupiter, and not remnants of a meteorite nebula, this calculation of the theoretical age of the Earth is just a rough estimate.

Shmidt further assumed that within the conglomerating meteorites, the matter got warm to such an extent that it became plastic from the impacts of meteorites in the process of the Earth's rapid growth and, most important, from the liberation of heat during radioactive processes. For this, temperatures of about 1,000 °C would have been completely sufficient. As the meteorite matter grew softer, the lighter stony masses floated up to the surface, and the heavier ferrous masses gradually sank below. In this way, the Earth gradually divided into a dense core and a lighter envelope, with some intermediate region in which malleable ferrous and stony masses did not separate completely.

It is believed now that the Earth's core is not ferrous, but silicate like the crust, in a heavily compressed metal-like state under high pressure of the overlying layers. In a layer where the pressure is 1,400,000 atmospheres, these properties of the silicate interiors of the Earth arise as a discontinuity. If we accept this point of view, the rising of light matter and sinking of heavy matter within the Earth proceeds slowly and is still far from completed.

The heating of the interior parts of the Earth is still continuing, and is due to the accumulation of heat given off by the radioactive decay within its matter.

The remnants of meteoric matter that did not become part of planets continued to rotate near the Sun and passing close to the planets were captured by them. In the flattened meteorite cloud which had formed around the planets, meteorites continued to collide and thus satellites were formed. Naturally, more massive planets, making more captures, could acquire a larger number of satellites.

As the majority of the meteorites used as construction material for satellites moved in the same forward direction near the Sun, and primarily close to the plane of ecliptic, the orbits of the satellites appeared to be close to this plane. The direction of their revolution turned out to be in accord with the movements common for all members of the solar system. Only in rare cases where in the distribution of speeds or densities of the meteor swarm a large asymmetry appeared, did planets and satellites arise with a retrograde rotation (Uranus with its satellites, Neptune's satellite, and the distant satellites of Jupiter and Saturn).

The rotation of the planets, which none of the former theories could explain satisfactorily, is explained by the theory of Shmidt as follows. It was caused by meteoric impacts, and at that in the same direction in which it revolves around the Sun. If in a planet's neighbourhood meteorites with orbits little elongated and little inclined to the average plane of the solar system were not sufficiently predominant, retrograde rotation might arise, thus explaining the only known case of this – the rotation of Uranus.

We will dwell for a while on a question that will not perhaps be as interesting for the reader as the preceding questions – angular momentum, which earlier theories could not explain.

We recall that the planets have the lion's share of the angular momentum in the solar system (i.e. the sum of the products of the masses of particles times their speed and times the distance from the centre of rotation). The Sun, with its slow rotation, accounts for a very small share of the angular momentum.

Shmidt's computation showed that the Sun, if initially it did not rotate, was to go into rotation by meteoritic impacts.

Shmidt succeeded in obtaining a formula that proves that the product

$$\frac{m^{2/3}\sqrt{R}P}{r^2} = w$$

should be constant or almost constant for all the planets. In this product, m is the mass of the planet, R is its distance from the Sun, r its radius, and P is the period of its rotation. As a matter of fact, this is how it turns out. Jupiter and Saturn display the greatest deviation from this law. But for a number of reasons, we were inclined to think even earlier that

the apparent radius of these planets, which is put in this formula, is not the actual radius of their solid surface – it is the radius of the apparent limit of their vast and dense atmosphere. In order to obtain w for Jupiter close to the one obtained for planets of the type of the Earth and Mars (which do not inspire such suspicions), it has to be assumed that for Jupiter the average density is the same as for the Earth, and that then it is only 6.8 times larger than the Earth (along the diameter). Almost half of its apparent radius in this case comprises the thickness of its vast nontransparent atmosphere. But Jeffreys earlier obtained about the same relation for the dimensions of the planet and its atmosphere although his arguments were completely different.

As regards Mercury and Venus, their initial rotation has been retarded to the present by the tidal action of the Sun.

In the same manner, but to a lesser degree, the Sun and the Moon retarded the daily rotation of the Earth by their tidal effect. Formerly, the Earth rotated more rapidly.

The merging of meteorites which moved along elongated ellipses with differently disposed semimajor axes leads to movement along an orbit that is more circular. The more meteorites conglomerate, i.e. the greater the variety of directions of the semimajor axes of their orbits, the closer to a circle will be the orbit of the planet. Actually, the orbits of the large planets – Jupiter and Saturn – are less elongated than those of Mercury and Mars.

But how are the planets distributed in their distances from the Sun? The answer to this question, found by Shmidt, turned out to be unexpectedly simple. It turns out that the angular momentum computed for a unit mass of a planet will increase in an arithmetic progression in going from one planet to the next. For bodies moving in circular orbits, the angular momentum (for a unit mass) is proportional to the square root of the orbit's radius. Consequently, the square roots of the distances of the planets from the Sun ($\sqrt{R}$) should increase in an arithmetic progression.

This law agrees perfectly with the actual distribution of the planets' distances from the Sun if we will only consider separately the group of planets distant from the Sun (from Jupiter to Pluto) and the group of planets close to the Sun (from Mercury to Mars). As said above, a part of the meteorites in the region of the planets of the second group fell in the Sun and, therefore, in considering their distances from the Sun, they should not be combined with the planets distant from the Sun. For planets close to the Sun, $\sqrt{R}$ increases on average by 0.20 in going from one planet to the next. Therefore, taking as the initial value of $\sqrt{R}$ its true value for Mercury, the following table can be prepared:

	Mercury	Venus	Earth	Mars
$\sqrt{R}$	0.62	$0.62 + 0.20 = 0.82$	$0.62 + 2 \cdot 0.20 = 1.02$	$0.62 + 3 \cdot 0.20 = 1.22$
R_{computed}	0.38	0.67	1.04	1.49
R_{true}	0.38	0.72	1.00	1.52

The first line shows the method of computing $\sqrt{R}$, the second line gives the computed value of the distances of the planets, and the third line gives the true distances. Good agreement is obtained.

For planets distant from the Sun, the average increase in $\sqrt{R}$ is equal to 1.00 and, therefore, taking as the initial value for $\sqrt{R}$ its true value for Jupiter, we obtain:

	Jupiter	Saturn	Uranus	Neptune	Pluto
$\sqrt{R}$	2.28	3.28	4.28	5.28	6.28
R_{computed}	5.20	10.8	18.3	27.9	39.4
R_{true}	5.20	9.5	19.2	30.1	39.5

Perfect agreement of the computed and true distance is obtained. Thus, it would seem that Shmidt succeeded in explaining the law of planet distances without having obtained any theoretical foundation in former cosmogonic theories. Several other cosmogonic theories of recent times also explain this phenomenon, but in different ways.

This is just one of the many cosmogonic hypotheses.

PART 1

Pictorial material for the present edition was kindly supplied by Urania-Verlag (GDR) – drawings (by Wolfgang Parschau); colour photographs pp. 6, 51, 53 (source: "Znaniye", Moscow), 52, 64 (Lothar Willmann), 2, 165, 169 (source: California Institute of Technology and Carnegie Institution of Washington), 74 (courtesy of NASA); black-and-white photographs pp. 14, 15, 19, 20, 21, 121, 130 (source: "Znaniye", Moscow), 43, 48, 49, 57, 65, 66, 68, 71, 144 (courtesy of NASA), 81 (source: Mount Wilson and Palomar Observatories, Pasadena), 113 (source: Sternwarte Sonneberg), 147, 162 (source: Sternwarte Pulsnitz), 114, 166 (source: Karl-Schwarzschild-Observatorium, Tautenburg), 151 (source: Museum für Naturkunde, Berlin, GDR), 127 (Dr Oleak), 67 (archives of W. A. Bronschten), 29, 30, 36, 46, 59, 63, 80, 156 (archives of Urania-Verlag);
Moscow Institute of Engineers for Geodesy, Aerial Photography and Cartography (USSR) – colour photographs pp. 38, 47, 50, 73.

PART 2

Pictorial material for the present edition was kindly supplied by Urania-Verlag (GDR) – drawings (by Wolfgang Parschau); colour photographs pp. 177 (Dr Staude), 205 (Dr Dorschner Dr Friedemann); black-and-white photographs pp. 194, 202, 212 (source: "Znaniye", Moscow), 241 (source: Mount Wilson and Palomar Observatories, Pasadena), 178, 185, 187 (source: Sonnenobservatorium Einsteinturm, Potsdam), 174, 248, 253, 254, 255, 258, 271, 276, 281, 282, 286, 294 left, 312, 313 (source: Karl-Schwarzschild-Observatorium, Tautenburg), 179, 180, 184, 243, 283, 289, 294 right (archives of Urania-Verlag)

TO THE READER

Mir Publishers would be grateful for your comments on the content, translation and design of this book.

We would also be pleased to receive any other suggestions you may wish to make.

Our address is:

Mir Publishers
2 Pervy Rizhsky Pereulok
I-110, GSP, Moscow, 129820
USSR

Printed in the Union of Soviet Socialist Republics

This book by a prominent Soviet astronomer and imaginative interpreter of the history and latest knowledge of astronomy, an age-old science that has come through two shattering decades of exciting discovery and research, can be dubbed a popular encyclopaedia of astronomy. It is a vivid and fascinating account of almost all the celestial bodies with which astronomy is concerned: planets, comets, meteors, the Sun and stars, gaseous and dust nebulae, quasars, black holes, etc. The tome is lavishly illustrated throughout with a mixture of photographs and specially commissioned artwork in full colour. The book will appeal to the general reader with no astronomical and mathematical background, fascinated by the wonders of the Space Age.